

明 細 書

言語モデル生成蓄積装置、音声認識装置、言語モデル生成方法および
音声認識方法

5

技術分野

本発明は、言語モデル生成蓄積装置、音声認識装置等に関し、特に、
統計的言語モデルを用いた音声認識装置および音声認識方法等に関する
ものである。

10

背景技術

近年、音声認識装置において、その性能を高めるために言語モデルを
用いる方法が研究されている。

言語モデルとしては、基本的な単語 2 グラム又は単語 3 グラムの単語
15 N グラムモデルが広く用いられている（例えば、非特許文献 1 参照）。

ここで、単語 N グラムを用いた言語尤度の計算について説明する。

まず、単語列 W_1, W_2, \dots, W_L の言語尤度 $\log P(W_1, W_2, \dots, W_L)$ は、条件付き確率を用いて下記に示される式（1）で表される。

$$\log P(W_1, W_2, \dots, W_L) = \sum_{i=1}^L \log P(W_i | W_1, W_2, \dots, W_{(i-1)}) \dots (1)$$

20

式（1）右辺の条件付き確率 $P(W_i | W_1, W_2, \dots, W_{(i-1)})$ は、先行単語列 $W_1, W_2, \dots, W_{(i-1)}$ の後に当該単語 W_i が生起する確率であり、この先行単語列を $N-1$ 個で近似するのが単語 N グラムモデルである。そして、先行単語列を 1 個で近似する単語 2 グラム

では、下記に示される式（２）の近似式で表される。

$$P(W_i | W_1, W_2, \dots, W(i-1)) \doteq P(W_i | W(i-1)) \quad \dots(2)$$

同様に、先行単語列を２個で近似する単語３グラムでは、下記に示される式（３）の近似式で表される。

$$P(W_i | W_1, W_2, \dots, W(i-1)) \doteq P(W_i | W(i-2), W(i-1)) \quad \dots(3)$$

5

音声認識の際には、認識途中の単語列候補に対して、単語音声の確率モデルであるHMM（Hidden Markov Model）等の音響モデルを用いて音響尤度を求め、さらに上記のようにして言語尤度を求め、それらを重み付き加算した総合尤度で単語列候補に序列をつけて、

10 音声認識を行う。

ところで、Nグラムモデルには多種のバリエーションがあるが、ここでは、特に本件と関連性のある３つの従来技術を以下に説明する。

Nグラムモデルの第１バリエーションとして、単語集団の中で、共通の性質を持つものをクラス化して確率を計算するものがある（例えば、
15 特許文献１参照）。なお、この技術を、以下「第１の従来例」とも記す。この第１の従来例のクラスNグラムモデルでは、単語Nグラムがクラスを用いて式（４）のように近似される（N＝２の場合）。ここで、C iはクラス化した単語を示す。

$$P(W_i | W(i-1)) \doteq P(C_i | C(i-1)) \times P(W_i | C_i) \quad \dots(4)$$

20 クラスを介して言語尤度を求めることで、学習データの少ない単語列

に対して言語尤度の精度が悪いというデータ量が不十分による問題に対して効果がある。

次に、Nグラムモデルの第2バリエーションとして、頻度の高い単語列を接続した新しい可変長単位を作り、可変長単位間のNグラムを用いるものがある（例えば、特許文献2参照）。なお、この技術を、以下「第2の従来例」とも記す。この第2の従来例のクラスNグラムモデルは、単位に関する工夫であり、式（2）および式（4）に従う。この第2の従来例によれば、1単語よりも長い単位を用いることで、より長い文脈を考慮した言語尤度を求めることができる。

さらに、Nグラムモデルの第3バリエーションとして、クラスNグラムにおける名前等の一部のクラスを単語ではなく、音節等のより細かい単位の列として、別のNグラムで表したものがある（例えば特許文献3参照）。なお、この技術を、以下「第3の従来例」とも記す。つまり、この第3の従来例は、式（4）の右辺第2項を下記に示される式（5）のように近似するものである。

$$P(W_i | C_i) \doteq P(P_1, P_2, \dots, P_j | C_i) \quad \dots(5)$$

ここで、 P_1, P_2, \dots, P_j は、単語の発音を表す単位の列を示す。

式（5）の右辺は、さらに、音節の並びに関する確率と、単語内の音節数に関する確率の積として表すことで、高精度化され、名前等の数が多いクラスを効率的に表すことができる。

以上の全ての場合について、確率値は、学習用のテキストコーパスを処理して、機械的に求めることができる。

図1は、第3の従来例による音声認識装置の機能構成を示すブロック図である。

図 1 に示されるように、音声認識装置 900 は、入力音声を取り込んで特徴パラメータを抽出する音響処理部 901 と、特徴パラメータと単語とを照合する音響尤度計算部 910 と、話者の音響的特徴をモデル化した音響モデル部 911 と、認識単語の発音を記述した単語辞書部 912 と、クラス N グラムおよびクラス依存音節 N グラムを参照して単語列仮説を生成する単語列仮説生成部 920 と、単語の文脈を評価するクラス N グラム蓄積部 9215 と、音節の文脈を評価するクラス依存音節 N グラム蓄積部 9223 と、認識対象となる文表現を多数蓄積した文表現コーパス蓄積部 9211 と、文表現を形態素解析する形態素解析部 9212 と、単語又は単語クラスの連鎖の統計量からクラス N グラムを生成するクラス N グラム生成部 9213 と、共通の性質を持つ単語をクラスとして定義した単語クラス定義部 9214 と、名前を蓄積した名前辞書部 9221 と、音節の連鎖の統計量を求めてクラス依存音節 N グラムを生成するクラス依存音節 N グラム生成部 9222 とから構成される。

次に、動作について説明する。この動作は、言語モデルを作成する事前処理と、入力された音声を逐次認識する認識処理とに大別される。

まず、言語モデルを作成する事前処理について説明する。

文表現コーパス蓄積部 9211 は、文表現を予め蓄積している。形態素解析部 9212 は、文表現コーパス蓄積部 9211 に蓄積されている文表現を形態素解析し、形態素である単語単位に分割する。クラス N グラム生成部 9213 は、単語クラス定義部 9214 を参照して、形態素解析済みのコーパス中の単語を単語クラスに置き換え、単語又は単語クラスの連鎖の統計量を求めてクラス N グラムを生成する。クラス N グラム蓄積部 9215 は、単語又は単語クラスの連鎖の統計量を蓄積する。

一方、名前辞書部 9221 は、名前の読みである音節列を予め蓄積する。クラス依存音節 N グラム生成部 9222 は、名前辞書部 9221 に

蓄積されている名前の読みである音節列から音節の連鎖の統計量を求めてクラス依存音節 N グラムを生成する。クラス依存音節 N グラム蓄積部 9 2 2 3 は、音節の連鎖の統計量を蓄積する。

次いで、入力された音声を逐次認識する認識処理について、以下に説明する。

入力された音声は、音響処理部 9 0 1 で処理され、特徴パラメータに変換される。特徴パラメータは、音響尤度計算部 9 1 0 により、音響モデル部 9 1 1 と単語辞書部 9 1 2 とを参照して、単語辞書内の各単語とマッチングされ、単語の発声区間と単語の音響尤度とからなる単語仮説群が出力される。単語仮説群は、単語列仮説生成部 9 2 0 において、単語列仮説に組み上げられ、クラス N グラムとクラス依存音節 N グラムとを参照して、式 (1) ~ 式 (5) により計算される言語尤度を加えられる。

このようにして、音響尤度と言語尤度の一次式により評価された尺度で序列をつけ単語列候補が認識結果となる。

【非特許文献 1】

大附、森、松岡、古井、白井：“新聞記事を用いた大語彙音声認識の検討” 信学技報、S P 9 5 - 9 0 (1 9 9 5 - 1 2)

【特許文献 1】

特開 2 0 0 0 - 2 5 9 1 7 5 号公報 (第 5 - 9 頁、第 1 図)

【特許文献 2】

特許第 3 0 0 4 2 5 4 号公報 (第 4 - 1 9 頁、第 1 図)

【特許文献 3】

特開 2 0 0 1 - 2 3 6 0 8 9 号公報 (第 4 - 1 1 頁、第 1 図)

音声認識装置は、認識精度を高くするための言語的推定精度の向上が

要求されている。

しかし、従来の方法は、テレビ番組や、映画のタイトル、例えば、「月
に向かって飛べ」や、「太陽を撃て」などのように、前後の単語に対して
は1単語としての第1の性質と、内部の表現としては複数の単語として
5 の第2の性質との二つの性質を有するものの処理については言語的推定
精度の向上が困難であるという問題があった。

すなわち、タイトル＝1単語として扱うと、単語の種類が多いために
認識辞書が大きくなってしまう。一方、タイトル＝単語列として扱うと、
タイトルの前後を含む文脈は2グラムや3グラムの範疇外であるため、
10 制約が緩慢になる。より具体的には、第1の従来例および第2の従来例
においては、単位を決めた後は、その単位の2、3個分の文脈を考慮す
るので、単位の長さに応じて制約が緩慢又は辞書の増加という問題のど
ちらかに直面する。また、第3の従来例においては、タイトルを前後の
単語に対しては1単語として扱うが、タイトル内の処理は発音列として
15 モデル化するという二重構造であるために、長いタイトルの発音に対す
る推定精度に限界がある。

そこで、本発明は、テレビ番組のタイトルのように1単語と複数単語
の二重の性質を持つものを扱え、言語尤度の推定精度と、認識辞書のコ
ンパクト性とを両立し、認識精度を高めることができる言語モデル生成
20 蓄積装置および音声認識装置等を提供することを目的とする。

発明の開示

上記目的を達成するために、本発明に係る言語モデル生成蓄積装置に
おいては、音声認識用の言語モデルを生成し、蓄積する言語モデル生成
25 蓄積装置であって、複数のテキストを、特定の言語的特性を有する単語
列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデル

を生成し、蓄積する上位Nグラム言語モデル生成蓄積手段と、前記単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルを生成し、蓄積する下位Nグラム言語モデル生成蓄積手段とを備えることを特徴とする。

- 5 このため、言語尤度を計算する際に、共通の性質を持つ単語列を単語列クラスとして扱い、入れ子構造を持ったNグラムを用いて、前後の単語に対しては上位階層のクラスNグラムにより1単位として扱い、クラス内は下位階層の単語Nグラムにより単語列として扱うことができる。これにより、長い文脈、単語列クラスを構成する単語列に対する言語尤度
10 の推定精度と、認識辞書のコンパクト性を両立した言語モデル生成蓄積装置ひいては音声認識装置を得ることができる。

- また、本発明に係る言語モデル生成蓄積装置においては、前記上位Nグラム言語モデル生成蓄積手段と、前記下位Nグラム言語モデル生成蓄積手段とは、異なるコーパスを用いて言語モデルをそれぞれ生成すること
15 を特徴とすることもできる。

- これにより、上位の言語モデルと下位の言語モデルをそれぞれ独立に構築することが可能となるため、コーパス収集が容易となる。また語彙の変化等による言語モデルの再構築が必要な場合においても、関連する上位又は下位の言語モデルのみの再構築でよいことになり、言語モデル
20 のメンテナンスが容易となるという効果を有する。

- また、前記下位Nグラム言語モデル生成蓄積手段は、下位Nグラム言語モデル用のコーパスを更新するコーパス更新手段を有し、下位Nグラム言語モデル生成蓄積手段は、更新されたコーパスに基づいて下位Nグラム言語モデルを更新生成することを特徴とする構成としてもよい。

- 25 これにより、コーパスに新番組のタイトルの自動登録等が可能となり、言語モデルのメンテナンスがさらに容易となるという効果を有する。

また、本発明に係る言語モデル生成蓄積装置においては、前記下位 N グラム言語モデル生成蓄積手段は、前記単語列クラスの単語の系列を意味を有する最小の言語単位である 1 以上の形態素に解析し、当該単語列クラスに依存して当該形態素の系列をモデル化することにより前記下位
5 N グラム言語モデルを生成することを特徴とする構成とすることもできる。

これにより、単語列クラスを構成する単語列からクラス依存単語 N グラムを求めることができるため、十分な学習データ量を確保でき、高い認識精度が得られるという作用を有する。

10 また、本発明に係る言語モデル生成蓄積装置においては、前記上位 N グラム言語モデル生成蓄積手段は、形態素に解析されたテキストに含まれる前記単語列クラスを仮想的な仮想単語に置き換え、当該仮想単語および各単語からなる系列をモデル化することにより前記上位 N グラム言語モデルを生成することを特徴としてもよい。

15 これにより、単語列クラスの定義を合わせた単語列クラスを含むテキストと、単語列クラスを構成する単語列との両方を考慮し、クラス N グラムを求めるため、高い認識精度が得られるという作用を有する。

また、本発明に係る言語モデル生成蓄積装置においては、前記下位 N グラム言語モデル生成蓄積手段は、前記単語列クラスに出現する単語のうち、特定の単語を単語列クラスの構成単語に含めない除外単語とする
20 か否かをその言語的特性に基づいて判断し、判断の結果前記除外単語について単語の読みに対応するモーラおよびモーラを結合した単位に切り分ける除外単語判断手段を有し、前記言語モデル生成蓄積装置は、さらに前記モーラおよびモーラを結合した単位の系列をモデル化して、確率
25 の対数値である言語尤度を前記単語列クラス又は該除外単語の言語的特性に依存して付与することによりクラス依存モーラ N グラムを生成し、

蓄積するクラス依存モーラNグラム生成蓄積手段を備えることを特徴とすることもできる。

これにより、単語列クラスに含まれる単語の一部をより細かな単位で表現することが可能であるため、音声認識辞書の語彙数の肥大化を避けてあらゆる種類の単語列クラスを高い精度で認識できるという作用を有する。

また、本発明に係る言語モデル生成蓄積装置においては、前記言語モデル生成蓄積装置は、さらに、テキストを形態素解析し、さらに構文解析し、予め決められた基準により選択された構文木上のノードに着目して、前記テキストを複数の階層構造とする構文木を生成する構文木生成手段を備え、前記上位Nグラム言語モデル生成蓄積手段は、前記着目したノードから上層を構成する第1の部分木を用いて構文木用の前記上位Nグラム言語モデルを生成し、前記下位Nグラム言語モデル生成蓄積手段は、前記着目したノードから下層を構成する第2の部分木を用いて生成された構文木用の下位Nグラム言語モデルを生成することを特徴としてもよい。

これにより、特定のノードに着目することで、構文木を容易に分割することができ、クラスNグラムによる長い文脈の評価と、クラス依存単語Nグラムによる単語列クラスを構成する単語列の評価との両方を考慮することで高い認識精度が得られるという作用を有する。

また、本発明に係る言語モデル生成蓄積装置においては、前記下位Nグラム言語モデル生成蓄積手段は、前記第2の層より下層の部分木に出現する特定の単語を、予め決められた言語的特性に基づいて部分木の構成単語に含めない除外単語と判断する言語モデル生成用除外単語判断手段を有し、前記下位Nグラム言語モデル生成蓄積手段は、前記除外単語については単語の読みに対応するモーラおよびモーラを結合した単位に

分解し、該モーラおよび結合モーラの系列を該除外単語の構文木上での位置および該除外単語の言語的特性に依存してモデル化することにより下位Nグラム言語モデルを生成することを特徴としてもよい。

また、本発明に係る言語モデル生成蓄積装置においては、前記言語モデル生成蓄積装置は、さらに、テキストを形態素解析し、さらに構文解析し、予め決められた基準により選択された構文木上のノードに着目して、前記テキストを複数の階層構造とする構文木を生成する構文木生成手段を備え、前記上位Nグラム言語モデル生成蓄積手段は、前記構文木の最上位層を構成する第1の部分木を用いて前記上位Nグラム言語モデルを生成し、前記下位Nグラム言語モデル生成蓄積手段は、第2より下位の層を構成する各部分木を、より上位の層における該部分木の位置付けに基づいて分類し、夫々の分類ごとの部分木を用いて前記下位Nグラム言語モデルを生成することを特徴としてもよい。

これにより、構文木生成部における構文解析結果に基づいて、大量のテキストから自動的にクラスNグラムとクラス依存単語Nグラムの両方を生成することができるという作用を有する。

また、本発明に係る言語モデル生成蓄積装置においては、前記下位Nグラム言語モデル生成蓄積手段は、前記第2の層より下層の部分木に出現する特定の単語を、予め決められた言語的特性に基づいて部分木の構成単語に含めない除外単語と判断する言語モデル生成用除外単語判断手段を有し、前記下位Nグラム言語モデル生成蓄積手段は、前記除外単語については単語の読みに対応するモーラおよびモーラを結合した単位に分解し、該モーラおよび結合モーラの系列を該除外単語の構文木上での位置および該除外単語の言語的特性に依存してモデル化することにより下位Nグラム言語モデルを生成することを特徴とすることもできる。

これにより、構文中の特定の関係に基づく一部の単語を、単語より小

さい単位で表現して認識することが可能となり、また、大量のテキストを構文解析した結果から、該構文中の特定の関係に基づいて、クラス依存モーラ N グラムを自動構築することができるという作用を有する。

また、本発明に係る言語モデル生成蓄積装置においては、前記上位 N
5 グラム言語モデル生成蓄積手段は、前記単語列クラスを含む長さ N の単語連鎖とその単語連鎖が生起する確率とを対応付けて上位 N グラム言語モデルを生成することを特徴としてもよい。

これにより、単語列クラスの生起確率による長い文脈の評価を考慮することで高い認識精度が得られるという作用を有する。

10 さらに、本発明に係る言語モデル生成蓄積装置においては、下位 N グラム言語モデル生成蓄積手段は、前記単語列クラスを構成する単語の長さ N の連鎖とその単語連鎖が生起する確率とを対応付けて下位 N グラム言語モデルを生成することを特徴としてもよい。

これにより、単語列クラスを構成する単語列の生起確率による単語列
15 クラスを構成する単語列の評価を考慮することで高い認識精度が得られるという作用を有する。

なお、本発明は、このような言語モデル生成蓄積装置として実現することができるだけでなく、このような言語モデル生成蓄積装置を備える音声認識装置として実現したり、言語モデル生成蓄積装置および音声認
20 識装置が備える特徴的な手段をステップとする言語モデル生成方法および音声認識方法として実現したり、それらのステップをコンピュータに実行させるプログラムとして実現したりすることもできる。そして、そのようなプログラムは、CD-ROM等の記録媒体やインターネット等の伝送媒体を介して配信することができるのは言うまでもない。

25 以上の説明から明らかなように、本発明に係る言語モデル生成蓄積装置および音声認識装置によれば、言語尤度を計算する際に、共通の性質

を持つ単語列を単語列クラスとして扱い、入れ子構造を持ったNグラムを用いて、前後の単語に対しては上位階層のクラスNグラムにより1単位として扱い、クラス内は下位階層の単語Nグラムにより単語列として扱うことが可能となり、これにより、長い文脈、単語列クラスを構成する単語列に対する言語尤度の推定精度と、認識辞書のコンパクト性を両立した言語モデル生成蓄積装置ひいては音声認識装置を得ることが可能になるという効果が奏される。

よって、本発明により、認識精度が高くなり、音声認識対応の家電製品が普及してきた今日における本願発明の実用的価値は極めて高い。

10

図面の簡単な説明

図1は、従来の音声認識装置を示す図である。

図2は、本発明第1の実施の形態による音声認識装置の構成を示す図である。

15 図3は、本発明第1の実施の形態によるクラスNグラム生成蓄積部の構成を示す図である。

図4は、文表現コーパス蓄積部111の構成例を示す図である。

図5は、クラスNグラム蓄積部114の構成例を示す図である。

20 図6は、本発明第1の実施の形態によるクラス依存単語Nグラム生成蓄積部の構成を示す図である。

図7は、クラスコーパス蓄積部121の構成例を示す図である。

図8は、クラス依存単語Nグラム蓄積部124の構成例を示す図である。

図9は、単語列クラス定義蓄積部126の構成例を示す図である。

25 図10は、音声認識処理の動作を示すフローチャートである。

図11は、単語列仮説生成部80における単語列仮説の評価を示す図

である。

図 1 2 は、第 2 の実施の形態による音声認識装置の構成を示す図である。

5 図 1 3 は、第 2 の実施の形態による構文木生成部の構成を示す図である。

図 1 4 は、第 2 の実施の形態による構文木用クラス N グラム生成蓄積部の構成を示す図である。

図 1 5 は、第 2 の実施の形態による構文木用クラス依存単語 N グラム生成蓄積部の構成を示す図である。

10 図 1 6 (a) は、第 2 の実施の形態による構文解析結果を示す図である。

図 1 6 (b) は、第 2 の実施の形態による構文木の分割を示す図である。

15 図 1 7 は、第 3 の実施の形態における音声認識装置の構成を示す図である。

図 1 8 は、第 3 の実施の形態によるクラス N グラム生成蓄積部の構成を示す図である。

図 1 9 は、第 3 の実施の形態によるクラス依存単語 N グラム生成蓄積部の構成を示す図である。

20 図 2 0 は、第 3 の実施の形態によるクラス依存モーラ N グラム生成蓄積部の構成を示す図である。

図 2 1 は、クラス依存モーラ N グラム蓄積部 3 3 2 の構成例を示す図である。

25 図 2 2 は、単語列仮説生成部 8 0 における単語列の評価の様子を示す図である。

図 2 3 は、第 4 の実施の形態によるクラス依存単語 N グラム生成蓄積

部の構成を示す図である。

発明を実施するための最良の形態

以下、本発明の実施の形態について、図面を参照しながら説明する。

5 (実施の形態 1)

図 2 は、本発明の実施の形態 1 に係る音声認識装置の構成を示す機能ブロック図である。

図 2 に示されるように、音声認識装置 1 は、言語モデル生成蓄積装置 10 と、入力音声を取り込んで特徴パラメータを抽出する音響処理部 40 と、特定又は不特定の話者の音響的特徴をモデル化した音響モデル部 60 と、認識単語の発音を記述した単語辞書部 70 と、音響モデルおよび単語辞書を参照して特徴パラメータおよび単語を照合する単語照合部 50 と、言語モデル生成蓄積装置 10 のクラス N グラムとクラス依存単語 N グラムとを参照して単語照合結果から単語列仮説を生成し、認識結果を得る単語列仮説生成部 80 とを備える。

10

15

言語モデル生成蓄積装置 10 は、言語的な確率の対数値である言語尤度を単語列クラスを含む文脈に付与するためのクラス N グラムを生成し、生成したクラス N グラムを蓄積するクラス N グラム生成蓄積部 11 と、言語的な確率の対数値である言語尤度を単語列クラス内の単語系列に付与するためのクラス依存単語 N グラムを生成し、生成したクラス依存単語 N グラムを蓄積するクラス依存単語 N グラム生成蓄積部 12 とから構成される。

20

次いで、音声認識動作について説明する。この音声認識動作は、言語モデルを作成する事前処理と、入力された音声を逐次認識する認識処理とに大別される。

25

まず、言語モデル生成蓄積装置 10 のクラス N グラム生成蓄積部 11

およびクラス依存単語 N グラム生成蓄積部 1 2 の構成を順次説明する。

5 なお、言語モデルは、単語および単語列クラスの系列を評価するクラス N グラムと、単語列クラスを構成する単語列を評価するクラス依存単語 N グラムとから構成され、音声認識処理を実行する前に予め作成される。

まず、クラス N グラムの生成について、図 3 を用いて詳細に説明する。

図 3 は、クラス N グラム生成蓄積部 1 1 の機能構成を示すブロック図である。

10 図 3 に示されるように、クラス N グラム生成蓄積部 1 1 は、認識対象となる文表現がテキストとして予め多数蓄積された文表現コーパス蓄積部 1 1 1 と、文表現を形態素解析する文表現用形態素解析部 1 1 2 と、単語列クラス定義を参照して形態素結果から、単語や単語列クラスの連鎖の統計量を求めてクラス N グラムを生成するクラス N グラム生成部 1 1 3 と、クラス N グラムを蓄積し、単語列仮説生成部 8 0 に出力するクラス N グラム蓄積部 1 1 4 とから構成される。

クラス N グラム生成蓄積部 1 1 の文表現コーパス蓄積部 1 1 1 は、認識対象となる文表現のデータライブラリを予め多数蓄積する。

20 より詳しくは、文表現コーパス蓄積部 1 1 1 は、図 4 に示されるように、「明日の天気予報を録画して」、「明日の太陽を撃てを録画」、「知床の神秘を見る」などの比較的長い文表現であるテキストを予め蓄積する。

文表現用形態素解析部 1 1 2 は、文表現コーパス蓄積部 1 1 1 が蓄積している「明日の天気予報を録画して」などの比較的長い文表現であるテキストから、意味を有する最小の言語単位である形態素を解析する。例えば、上記文表現の「明日の天気予報を録画して」は、形態素解析に
25 より「＜ S S ＞－明日－の－天気－予報－を－録画－して－＜ S E ＞」となる。同様に、「明日の太陽を撃てを録画」、「知床の神秘を見る」は、

「＜ＳＳ＞－明日－の－太陽－を－撃て－を－録画－＜ＳＥ＞」、 「＜Ｓ
Ｓ＞－知床－の－神秘－を－見る－＜ＳＥ＞」となる。ここで、＜ＳＳ
＞、＜ＳＥ＞はそれぞれ、文頭、文末を意味する仮想的な単語である。

次に、クラスＮグラム生成部１１３は、形態素に解析されたテキスト
5 に含まれる単語列を抽出し、後述するクラス依存単語Ｎグラム生成蓄積
部１２から入力される単語列クラスを参照し、該当する単語列クラスが
存在する場合は、テキストに含まれる単語列クラスを仮想的な単語に置
き換え、単語又は単語列クラスの連鎖の統計量を求めることで単語又は
単語列クラスの連鎖とその確率とを対応付けたクラスＮグラムを生成す
10 る。形態素単位に分割された文表現は、クラスＮグラム生成部１１３に
おいて、単語列クラス定義を参照して、単語列クラス定義にある単語列
はクラスを示す仮想的な単語に置き換えた上で、１～Ｎ個の単語連鎖ご
とに頻度を計測し、確率モデルが生成される。このクラスを単語列クラ
スと呼ぶ。クラスＮグラム生成部１１３によって生成されたクラスＮグ
15 ラムは、クラスＮグラム蓄積部１１４に蓄積される。

例えば、＜ｔｉｔｌｅ＞という単語列クラスに「天気－予報」が定義
されている場合は、形態素解析結果は、「＜ＳＳ＞－明日－の－＜ｔｉ
ｔ
1 ｌ
e＞－を－録画－して－＜ＳＥ＞」と置き換えられる。同様に、＜ｔ
i
t
l
e＞という単語列クラスに「太陽－を－撃て」、「知床－の－神秘」
20 が定義されている場合は、形態素解析結果は、＜ＳＳ＞－明日－の－＜
t
i
t
l
e＞－を－録画＜ＳＥ＞、＜ＳＳ＞－＜ｔｉｔｌｅ＞－を－見
る－＜ＳＥ＞にそれぞれ置き換えられる。また、条件付き確率は、単語
３グラムモデルの場合、 W_3 が W_1 － W_2 の連鎖に後続する確率は、 W
1－ W_2 － W_3 の３つの組の連鎖の頻度を、 W_1 － W_2 の２つの組の連
25 鎖の頻度で割り算する $P(W_3 | W_1, W_2) = (W_1, W_2, W_3 \text{ の連鎖頻度}) / (W_1, W_2 \text{ の連鎖頻度})$ で求められ、単語２グラムモデ

ルの場合も同様に、 $P(W2 | W1) = (W1, W2 \text{ の連鎖頻度}) / (W1 \text{ の頻度})$ で求められる。

より詳しくは、単語 2 グラムモデルの場合、クラス N グラム生成部 1 1 3 は、 $\langle SS \rangle$ - 明日、明日 - の、の - $\langle title \rangle$ 、 $\langle title$
5 $e \rangle$ - を、を - 録画、録画 - して、して - $\langle SE \rangle$ 、 $\langle SS \rangle$ - 明日、
明日 - の、の - $\langle title \rangle$ 、 $\langle title \rangle$ - を、を - 録画、録画
- $\langle SE \rangle$ 、 $\langle SS \rangle$ - $\langle title \rangle$ 、 $\langle title \rangle$ - を、を - 見る、見る - $\langle SE \rangle$ 、…の頻度をそれぞれ求め、 $(W1, W2 \text{ の連鎖頻度})$
10 $/ (W1 \text{ の頻度})$ を計算することにより、それぞれの確率 $P(W2 | W$
1) を求める。そして、クラス N グラム生成部 1 1 3 は、図 5 に示され
るように単語の連鎖とその確率を対応付けてクラス N グラム蓄積部 1 1
4 に蓄積する。

これより、各単語連鎖の頻度を計測しておくことで、条件付き確率を計算でき、また、単語列クラスは単語と同様に扱うことができ、単語ごと
15 に条件付き確率が付加された言語モデルとなる。結果として、クラス N
グラムの役割は、「 $\langle SS \rangle$ - 明日 - の - $\langle title \rangle$ - を - 録画 - し
て - $\langle SE \rangle$ 」と置き換えることにより、各単語に条件付き確率を付加
することができる。

次に、クラス依存単語 N グラムの生成について、図 6 を用いて詳細に
20 説明する。

図 6 は、クラス依存単語 N グラム生成蓄積部 1 2 の機能構成を示すブロック図である。

図 6 に示されるように、クラス依存単語 N グラム生成蓄積部 1 2 は、
クラスコーパス蓄積部 1 2 1 と、クラス用形態素解析部 1 2 2 と、クラス
25 依存単語 N グラム生成部 1 2 3 と、クラス依存単語 N グラム蓄積部 1
2 4 と、単語列クラス定義生成部 1 2 5 と、単語列クラス定義蓄積部 1

26 とから構成される。

クラスコーパス蓄積部 121 は、意味的な性質や、構文的な性質が同一である単語列（例えば、テレビ番組のタイトルや、人名等）のデータライブラリを予め蓄積する。

- 5 より詳しくは、クラスコーパス蓄積部 121 は、図 7 に示されるように、「天気予報」、「太陽を撃て」、「知床の神秘」などのタイトルや、「チャーリー海」、「池乃キングヨ」等の単語列を予め蓄積する。このような単語列は、例えばこれから放送される番組表に基づいて予め入力される。

クラス用形態素解析部 122 は、クラスコーパスを形態素解析する。

- 10 具体的には、クラス用形態素解析部 122 は、クラスコーパス蓄積部 121 が蓄積している「天気予報」の様なテレビ番組名などの比較的短く、共通の性質を持つ単語列を、形態素単位に解析する。例えば、単語列「天気予報」は、形態素解析により「＜CS＞－天気－予報－＜CE＞」となる。ここで、＜CS＞、＜CE＞はそれぞれ、単語列クラスの始め、
15 単語列クラスの終わりを表す仮想的な単語である。

- クラス依存単語 N グラム生成部 123 は、形態素解析結果を処理し、単語の連鎖の統計量を求めて、単語列とその確率とを対応付けた情報であるクラス依存単語 N グラムを生成する。つまり、クラス依存単語 N グラム生成部 123 は、入力された形態素単位の単語列における単語連鎖
20 の頻度を計測し、確率モデル化し、クラス依存単語 N グラムを生成し、生成したクラス依存単語 N グラムをクラス依存単語 N グラム蓄積部 124 に蓄積する。

- より詳しくは、単語 2 グラムモデルの場合、クラス依存単語 N グラム生成部 123 は、タイトルについて、＜CS＞－天気、天気－予報、予
25 報－＜CE＞、＜CS＞－太陽、太陽－を、を－撃て、撃て－＜CE＞、＜CS＞－知床、知床－の、の－神秘、神秘－＜CE＞、…の頻度をそ

れぞれ求め、 $(W1, W2 \text{ の連鎖頻度}) / (W1 \text{ の頻度})$ を計算することにより、それぞれの確率 $P(W2 | W1)$ を求める。人名の場合についても、同様である。そして、クラス依存単語 N グラム生成部 1 2 3 は、図 8 に示されるように単語列とその確率を対応付けてクラス依存単語 N
5 グラム蓄積部 1 2 4 に蓄積する。この結果、クラス依存単語 N グラム生成部 1 2 3 により、形態素に分割された単語列は、単語連鎖の頻度をクラス N グラムと同様に計測することで、確率モデル化された言語モデルとなる。

クラス依存単語 N グラム蓄積部 1 2 4 は、クラス依存単語 N グラム生成部 1 2 3 が生成したクラス依存単語 N グラムを蓄積する。このクラス
10 依存単語 N グラム蓄積部 1 2 4 に蓄積されたクラス依存単語 N グラムは、音声認識の際に単語列仮説生成部 8 0 に参照される。

単語列クラス定義生成部 1 2 5 は、クラスコーパスの形態素解析結果から共通の性質を持つ単語列をクラスとして定義した単語列クラスの定義を生成する。具体的には、単語列クラス定義生成部 1 2 5 は、形態素
15 単位に解析された単語列から、共通の性質を持つ単語列をクラスとする単語列クラスの定義を生成する。ここで、単語列クラスには、`title` を集めた単語列であるコーパスに「天気予報」、「太陽を撃て」等があり、「天気ー予報」、「太陽ーを一撃て」などの単語列は `<title>` クラスであると定義する。
20

単語列クラス定義蓄積部 1 2 6 は、単語列クラス定義生成部 1 2 5 が生成した単語列クラス定義を蓄積する。この単語列クラス定義は、上記クラス N グラムの生成の際にクラス N グラム生成蓄積部 1 1 のクラス N
グラム生成部 1 1 3 によって参照される。

つまり、単語列クラス定義生成部 1 2 5 は、`<CS>` - 天気、天気ー
25 予報、予報 - `<CE>`、`<CS>` - 太陽、太陽ーを、を - 撃て、撃て -

<CE>、<CS>—知床、知床—の、の—神秘、神秘—<CE>、…
については、「t i t l e」であると定義し、<CS>—チャーリー—海
—<CE>、<CS>—池乃—キングョ—<CE>…については、人名
であると定義する。そして、単語列クラス定義生成部 1 2 5 は、図 9 に
5 示されるように単語列とその単語列クラスを対応付けて単語列クラス定
義蓄積部 1 2 6 に蓄積する。これによって、クラス N グラム生成部 1 1
3 は、適切な単語列クラスを取得することができる。

次に、入力された音声を逐次認識する認識処理について説明する。

図 1 0 は、音声認識処理の動作を示すフローチャートである。

- 10 音響処理部 4 0 は、マイクロフォン等から入力された音声を取得する
と (S 1 1)、その音声を特徴パラメータに変換する (S 1 2)。ここで、
特徴パラメータには、線形予測分析から得られる L P C ケプストラムや、
M F C C (M e l F i l t e r e d C e p s t r u m C o e f f
i c i e n t) がある。特徴パラメータに変換されると、単語照合部 5
15 0 は、音響モデル部 6 0 および単語辞書部 7 0 を参照して、単語辞書内
の各単語とマッチングし、単語の発声区間と単語の音響尤度とからなる
単語仮説群に変換する (S 1 3)。ここで、音響モデルには単語音声の確
率モデルである H M M (H i d d e n M a r k o v M o d e l) な
どがあり、入力音声の特徴パラメータが音節等の音響的な単位である音
20 響尤度を与える。また、マッチングのアルゴリズムにはビタビアルゴリ
ズムなどがある。

- 次に、単語列仮説生成部 8 0 は、全ての単語仮説群について (S 1 4)、
単語区間を考慮して接続された単語列仮説に組み上げられ、クラス N グ
ラムおよびクラス依存単語 N グラムを参照して、後述する言語尤度を付
25 与する。このようにして、単語照合部 5 0 は、音響尤度および単語列仮
説生成部 8 0 による言語尤度から求まる一次式により評価された尺度

(得点)を用いて、序列をつけた単語列候補を評価する(S15, 16)。
より詳しくは、ある単語列仮説が a, b, c, d である場合、単語列仮
説生成部 80 は、図 11 に示されるように、クラスを含まない単語列 <
SS> - a - b - c - d - <SE> の確率 P(a, b, c, d)、a がク
5 ラス C であるとした単語列 <SS> - C - b - c - d - <SE> の確率
P(C, b, c, d) · P(a | C)、a, b がクラス C であるとした単
語列 P(C, c, d) · P(a, b | C)、…、a, b, c, d がクラス
C であるとした単語列 <SS> - C - <SE> の確率 P(a, b, c,
d | C) を総当たりの評価する。そして、単語列仮説生成部 80 は、
10 音声認識結果として得点の最大値 max を選び出し、音声認識処理を終
了する。

なお、この実施の形態 1 では単語照合が終わってから単語列仮説の生
成を行うように説明をしたが、単語照合と単語列仮説の生成とを同時進
行的に進めることが可能である。

15 次いで、言語尤度の計算方法について述べる。

なお、ここでは先行単語を 1 単語用いる場合で説明するが、先行単語
を 2 単語用いる場合についても同様に実施可能である。

まず、任意の単語列 W1, W2, …, WL の言語尤度は、下記に示さ
れる式 (6) によって計算する。

$$\log P(W1, W2, \dots, WL) \doteq \sum_{i=1}^L \log P\{Wi | W(i-1)\} \quad \dots(6)$$

20

上式 (6) 右辺の確率は、下記に示される式 (7) で求める。

$$P(W_i | W(i-1)) = \begin{cases} P1(W_i | W(i-1)) & \text{どちらも一般単語の場合} \\ P1(C_i | W(i-1)) \times P(W_i | C_i) & W_i \text{のみクラス単語の場合} \\ P2(W_i | W(i-1)) & \text{どちらもクラス単語の場合} \\ P2(CS | W(i-1)) \times P1(W_i | C(i-1)) & W(i-1) \text{のみクラス単語の場合} \end{cases} \dots(7)$$

ここで、P1はクラスNグラムによる確率、P2はクラス依存単語Nグラムによる確率である。また、共通の性質を持つ単語列に対して同じクラス記号を与えたものである単語列クラスに含まれる単語をクラス単語、それ以外の単語を一般単語とした。ただし、一般的には、特定の単語が、クラス単語であるか一般単語であるかの判断が困難であるため、式(7)の4つの確率を足したものを左辺の値としてもよい。

以上のように求められた言語尤度が、組み上げられた単語列仮説に付加され、単語列候補に序列をつけ、認識結果として出力される。

10 以下に、テレビ番組名である「太陽を撃て」を録画する場合の「明日の太陽を撃てを録画」という発声を例に、従来の計算例と、本発明の計算式とを示すことで違いを明確にするとともに、本発明の効果を示す。

まず、例文を単語の列に分割する3つの方法を説明する。

15 第1に、テレビ番組名を1単語として扱う場合の「明日ーのー太陽を撃てーをー録画」と、

第2に、テレビ番組名を3単語に分割して扱う場合の「明日ーのー太陽ーをー撃てーをー録画」とがあり、

まず、単語2グラムモデルにより、

第1の場合について、式(8)により計算する。

$$P(<SS> \text{-} \text{明日} \text{-} \text{の} \text{-} \text{太陽を撃て} \text{-} \text{を} \text{-} \text{録画} \text{-} <SE>)$$

\doteq

$$P(\text{明日} \mid <SS>) \times P(\text{の} \mid \text{明日}) \times P(\text{太陽を撃て} \mid \text{の}) \times P(\text{を} \mid \text{太陽を撃て}) \\ \times P(\text{録画} \mid \text{を}) \times P(<SE> \mid \text{録画})$$

…(8)

このモデルでは、「太陽を撃て」と同様に複数の単語の組み合わせからなるテレビ番組名の数が多くなり、認識辞書が大きくなってしまう。

次に、第2の場合の単語2グラムモデルについて、式(9)により計算する。

$$P(<SS> \text{-} \text{明日} \text{-} \text{の} \text{-} \text{太陽} \text{-} \text{を} \text{-} \text{撃て} \text{-} \text{を} \text{-} \text{録画} \text{-} <SE>)$$

\doteq

$$P(\text{明日} \mid <SS>) \times P(\text{の} \mid \text{明日}) \times P(\text{太陽} \mid \text{の}) \times P(\text{を} \mid \text{太陽}) \times \\ P(\text{撃て} \mid \text{を}) \times P(\text{を} \mid \text{撃て}) \times P(\text{録画} \mid \text{を}) \times P(<SE> \mid \text{録画})$$

…(9)

これらの各確率は、テレビ番組名を含んだ文表現コーパス蓄積部111から学習することになるが、学習データを十分に準備することは困難であるから、データ量が不十分な学習データとなり、一部の単語系列に対して音響的な確率ではなく、言語的な確率、単語の連鎖に関する確率の精度が劣化する。

式(9)では特に、テレビ番組名と前の単語の文脈P(太陽 | の)、テレビ番組名と後の単語の文脈P(を | 撃て)、テレビ番組名の中の文脈P(を | 太陽)、P(撃て | を)の確率は信頼性が低い。

このようなデータ量が不十分による問題に対処するために、単語をクラス化して用いることができる。

第1の場合について、テレビ番組名部分をクラスとして扱うと、下記

に示される式（１０）が得られる。

$$P(<SS> - \text{明日} - \text{の} - \text{太陽を撃て} - \text{を} - \text{録画} - <SE>)$$

\doteq

$$\begin{aligned} &P(\text{明日} | <SS>) \times P(\text{の} | \text{明日}) \times P(<title> | \text{の}) \times \\ &P(\text{太陽を撃て} | <title>) \times P(\text{を} | <title>) \times P(\text{録画} | \text{を}) \times \\ &P(<SE> | \text{録画}) \end{aligned} \quad \dots(10)$$

- これは、テレビ番組名の前後の文脈が、 $P(<title> | \text{の})$ と $P(\text{を} | <title>)$ で表されるために、データ量が不十分による問題に対処できるが、やはり「太陽を撃て」と同様のテレビ番組名の数が多くなり、認識辞書が大きくなってしまう。

さらに、第３の方法として、従来例３を用いると、下記に示される式（１１）が得られる。

$$P(<SS> - \text{明日} - \text{の} - \text{太陽を撃て} - \text{を} - \text{録画} - <SE>)$$

\doteq

$$\begin{aligned} &P(\text{明日} | <SS>) \times P(\text{の} | \text{明日}) \times P(<title> | \text{の}) \times P(\text{タ} | <TS>) \times \\ &P(\text{イ} | \text{タ}) \times P(\text{ヨ} | \text{イ}) \times P(\text{ウ} | \text{ヨ}) \times P(\text{ヲ} | \text{ウ}) \times P(\text{ウ} | \text{ヲ}) \times P(\text{テ} | \text{ウ}) \\ &P(<TE> | \text{テ}) \times P(\text{を} | <title>) \times P(\text{録画} | \text{を}) \times P(<SE> | \text{録画}) \end{aligned} \quad \dots(11)$$

- これは、テレビ番組名の前後の文脈が、 $P(<title> | \text{の})$ と $P(\text{を} | <title>)$ で表されるために、データ量が不十分による問題に対処できて、かつ、テレビ番組名が音節列で表されるために認識辞書も小さい。

- しかし、テレビ番組名を音節列で表すために、制約がゆるく、認識精度が得られない。特に、テレビ番組名が長い場合に、全ての音節を正しく認識することは困難である。

従来例 3 は数音節を 1 単位として用いることもできるが、形態素である単語等の単位であれば、意味的な役割や構文的な役割と結びつくが、発音を表す音節の列は、結びつかなかったり、同音異義語が共有されたりするなどの問題がある。

- 5 これらに対して、本発明の実施の形態 1 では、下記に示される式 (1 2) として計算する。

$P(<SS>-明日-の-太陽を撃て-を-録画-<SE>)$

\doteq

$P(明日 | <SS>) \times P(の | 明日) \times P(<title> | の) \times P(太陽 | <CS>) \times$
 $\times P(を | 太陽) \times P(撃て | を) \times P(<CE> | 撃て) \times P(を | <title>) \times$
 $P(録画 | を) \times P(<SE> | 録画)$

...(12)

- これは、テレビ番組名の前後の文脈が、 $P(<title> | の)$ と、 $P(を | <title>)$ とで表されるために、データ量が不十分による問題に対処できて、かつ、テレビ番組名が形態素列で表されるために
 10 認識辞書 (クラス N グラム蓄積部 1 1 4 およびクラス依存単語 N グラム蓄積部 1 2 4) も小さい。しかも、テレビ番組名を形態素列で表すために、音節列で表す場合と比較して高い認識性能を確保できる。

- さらに、テレビ番組名部分の確率は、他の部分と比較して確率が低く、
 15 認識され難いという問題に対しては、クラス N グラムによる確率の代表的な値と、クラス依存単語 N グラムによる確率の代表的な値の差分を、クラス依存単語 N グラムによる確率に対してオフセットとして加え、発声区間全体において音声認識の尤度計算終了後、オフセット分を差し引くことで、他の音声認識結果候補の尤度との調節ができ、認識精度が上
 20 昇する。

(実施の形態 2)

図 1 2 は、本発明の実施の形態 2 に係る音声認識装置の機能構成を示すブロック図である。なお、実施の形態 1 の言語モデル生成蓄積装置 1 0 および音声認識装置 1 の構成と対応する部分に同じ番号を付し、その説明を省略する。

- 5 図 1 2 に示されるように、音声認識装置 2 は、上記した音声認識装置 1 の言語モデル生成蓄積装置 1 0 に代えて用いられる言語モデル生成蓄積装置 2 0 と、音響処理部 4 0 と、単語照合部 5 0 と、音響モデル部 6 0 と、単語辞書部 7 0 と、単語列仮説生成部 8 0 とを備える。

- 言語モデル生成蓄積装置 2 0 は、言語モデルを作成する事前処理において、構文解析アルゴリズムを用いて文表現の構文を解析することにより、クラス N グラムおよびクラス依存単語 N グラムを作成するものであり、テキストである文表現を構文解析し、文章の構造を階層的に表した構文木とする構文木生成部 2 1 と、入力された文表現から、クラス N グラムを生成、蓄積する構文木用クラス N グラム生成蓄積部 2 2 と、入力された文表現から、クラス依存単語 N グラムを生成、蓄積する構文木用クラス依存単語 N グラム生成蓄積部 2 3 とを備えて構成される。なお、
10 構文木用クラス N グラム生成蓄積部 2 2 と、構文木用クラス依存単語 N グラム生成蓄積部 2 3 とは、単語列仮説生成部 8 0 の要求に応じて、クラス N グラムおよびクラス依存単語 N グラムを単語列仮説生成部 8 0 に
15 出力する。

次に、構文木生成部 2 1 について詳細に説明する。

図 1 3 は、構文木生成部 2 1 の機能構成を示すブロック図である。

- 図 1 3 に示されるように、構文木生成部 2 1 は、上述した文表現コーパス蓄積部 1 1 1、文表現用形態素解析部 1 1 2 の他、構文解析部 2 1
20 1 と、構文木分割部 2 1 2 とを備えて構成される。

構文解析部 2 1 1 は、形態素解析された文の構文を解析する。

構文木分割部 2 1 2 は、構文木中のノードを選択するノード選択部を示し、選択されたノードから上層を構成する第 1 の部分木と下層を構成する第 2 の部分木とに構文木を分割する。

例えば、文表現コーパス蓄積部 1 1 1 に「彼は駅まで歩くと行った」
5 が蓄積されている場合、文表現用形態素解析部 1 1 2 は、「彼ーはー駅ー
までー歩くーとー言っーた」と解析する。構文解析部 2 1 1 は、C Y K
法などの公知の構文解析アルゴリズムにより解析し、図 1 6 (a) に示
されるように、文章の構造を表した構文解析結果である構文木を取得す
る。なお、図 1 6 (a) において、S 8 0 1 は文、S S 8 0 7 はサブ文、
10 P P 8 0 2 は後置句、V P 8 0 3 は動詞句、N P 8 0 4 は名詞句、P 8
0 5 は後置詞、V 8 0 8 は動詞、N 8 0 6 は名詞、T 8 0 9 は時制を表
す。

ここで、構文木分割部 2 1 2 においてノード「S S 8 0 7」を選択す
るように予め設定されており、構文木分割部 2 1 2 は、ノード「S S 8
15 0 7」に相当する部分を仮想的な単語として、「S S」と置き換え、図 1
6 (b) に示されるように、2 階層の構文木に変換する。なお、図 1 6
(b) において、8 1 0 は選択された S S ノードから上層を構成する第
1 の部分木を示し、8 1 1 は選択された S S ノードから下層を構成する
第 2 の部分木を示す。

20 次に、構文木分割部 2 1 2 は、第 1 の部分木 8 1 0 である「彼ーはー
S Sーとー言っーた」を構文木用クラス N グラム生成蓄積部 2 2 に出力
し、第 2 の部分木 8 1 1 である「駅ーまでー歩く」を構文木用クラス依
存単語 N グラム生成蓄積部 2 3 に出力する。

次いで、構文木用クラス N グラム生成蓄積部 2 2 について詳細に説明
25 する。

図 1 4 は、構文木用クラス N グラム生成蓄積部 2 2 の機能構成を示す

ブロック図である。

図 1 4 に示されるように構文木用クラス N グラム生成蓄積部 2 2 は、構文木用クラス N グラム生成部 2 2 1 と、構文木用クラス N グラム蓄積部 2 2 2 とから構成される。

5 構文木用クラス N グラム生成部 2 2 1 は、構文木分割部 2 1 2 が単語とみなした「S S」を含む各単語に条件付き確率を付加することによりクラス N グラムを生成する。構文木用クラス N グラム蓄積部 2 2 2 は、構文木用クラス N グラム生成部 2 2 1 が生成したクラス N グラムを蓄積する。

10 次いで構文木用クラス依存単語 N グラム生成蓄積部 2 3 について説明する。

図 1 5 は、構文木用クラス依存単語 N グラム生成蓄積部 2 3 の機能構成を示すブロック図である。

図 1 5 に示されるように、構文木用クラス依存単語 N グラム生成蓄積部 2 3 は、構文木用クラス依存単語 N グラム生成部 2 3 1 と、構文木用クラス依存単語 N グラム蓄積部 2 3 2 とから構成される。

構文木用クラス依存単語 N グラム生成部 2 3 1 は、構文木分割部 2 1 2 が単語とみなした「S S」を構成する単語に条件付き確率を付加することによりクラス依存単語 N グラムを生成する。構文木用クラス依存単語 N グラム蓄積部 2 3 2 は、構文木用クラス依存単語 N グラム生成部 2 3 1 が生成したクラス依存単語 N グラムを蓄積する。

このようにして得られるクラス N グラムとクラス依存単語 N グラムとにより、実施の形態 1 の場合と同様に、S S を含む長い文脈と、S S の中の短い文脈とを同時に扱うことができる。しかも、構文木分割部 2 1 2 で S S の中の短い文脈を分割するので、実施の形態 1 の場合のように
25 クラスコーパス蓄積部 1 2 1 を用意する必要がなくなる。

なお、図 16 に示した 2 階層の「基本的な単語 N グラム」に入れ子構造を導入した例で説明したが、従来の他の N グラムのバリエーションとの組み合わせも実施可能である。

例えば、タイトルクラス内を表す単語 N グラムを、似た性質を持つ単語集合でクラス化したクラス N グラムで表し、よくある単語連鎖を接続した可変長単位で表すことも可能である。

また、上位階層および下位階層の 2 階層に限らず、より多くの階層や再起的な入れ子構造、例えば、「彼は駅まで歩いたと思ったと言った」は、「彼は「駅まで歩いた」と思ったと言った」と分割することも可能である。

また、クラス N グラムとクラス依存単語 N グラムとを分けずに、共通の一つの言語モデルで表すことも可能である。

(実施の形態 3)

図 17 は、本発明の実施の形態 3 に係る音声認識装置の機能構成を示すブロック図である。なお、図 2 と同じ番号を付しているブロックの認識処理は、実施の形態 1 の音声認識装置 1 と同様の動作であるのため、ここでの説明は省略する。

図 17 に示されるように、音声認識装置 3 は、音響処理部 40 や、単語照合部 50、音響モデル部 60、単語辞書部 70、単語列仮説生成部 80 の他、言語モデル装置 30 および単語が単語列クラスの構成単語であるか否かを判定する認識用除外単語判定部 90 を備えて構成される。

認識用除外単語判定部 90 は、単語列クラスの生起確率に基づく言語尤度の計算において、クラス依存単語 N グラムのみによる計算を行うか、クラス依存モーラ N グラムも参照して計算を行うかを判定するものである。

言語モデル装置 30 は、クラス N グラムを生成し、生成したクラス N

グラムを蓄積するクラスNグラム生成蓄積部31と、クラス依存単語N
グラムを生成し、生成したクラス依存単語Nグラムを蓄積するクラス依
存単語Nグラム生成蓄積部32と、クラス依存モーラNグラムを生成し、
生成したクラス依存モーラNグラムを蓄積するクラス依存モーラNグラ
ム生成蓄積部33とから構成される。

実施の形態3による音声認識装置3は、音声認識装置1と同様に、言
語モデルを作成する事前処理と、入力された音声を逐次認識する認識処
理とに大別される。

次に、言語モデルを作成する事前処理について説明する。

10 言語モデルは、単語列クラスを含む文脈であるテキストを評価するク
ラスNグラムと、単語列クラスを構成する単語列を処理するクラス依存
単語Nグラムおよびクラス依存モーラNグラムとがあり、認識処理を実
行する前に作成する。

まず、クラスNグラムの生成について、詳細に説明する。

15 図18は、クラスNグラム生成蓄積部31の機能構成を示すブロック
図である。なお、図18において図3と同じ番号の付されているブロッ
クは、実施の形態1に示したものと同様のものである。

図18に示されるように、クラスNグラム生成蓄積部31は、文表現
コーパス蓄積部111や、文表現用形態素解析部112の他、クラス依
存単語Nグラム生成蓄積部32において予め取得された単語列クラス定
義を参照して、形態素解析結果から単語列クラスおよび一般の単語の属
するクラスの連鎖の統計量を求めて、クラス連鎖確率を求めるクラス連
鎖モデル生成部311と、単語のクラスから各々の単語が出力される確
率を求める単語出力モデル生成部312と、クラス連鎖モデル生成部3
11によるモデルと、単語出力モデル生成部312によるモデルを合わ
せてクラスNグラムとして蓄積するクラスNグラム蓄積部313とから

構成される。

クラスNグラム生成蓄積部31における処理は、実施の形態1において図3に示したものと同様であり、文表現コーパス蓄積部111が蓄積している「明日の天気予報を録画して」などの比較的長い文表現である
5 テキストを、文表現用形態素解析部112に入力し、テキストから意味を有する最小の言語単位である形態素を解析し、この結果をクラス連鎖モデル生成部311および単語出力モデル生成部312に出力する。

クラス連鎖モデル生成部311は、後述するクラス依存単語Nグラム生成蓄積部32に蓄積された単語列クラスに属する単語列が存在する場合
10 には、これを単語列クラスを示す仮想的なシンボルに変換し、また、そうでない通常の単語については、各単語の属するクラスを示すシンボルに変換する。このようにして得られたクラスシンボルの系列に対して、これらの連鎖の統計量を求めることで、クラス連鎖モデルを生成する。

また、単語出力モデル生成部312は、形態素解析結果の単語列から、
15 単語列クラスに属する単語を除いた全ての単語について、単語の出現数と、それぞれの単語の属するクラスの出現数の統計量を求め、クラスに対して単語が出現する確率を求めて、単語出力モデルとする。

これら2つのモデルは、クラスNグラム蓄積部313に格納されて、単語列仮説生成部80において、式(13)に示した言語尤度の計算に
20 おいて参照される。

例えば、文表現の「明日の天気予報を録画して」は、形態素解析により「<SS>—明日—の—天気—予報—を—録画—して—<SE>」となる。今仮に、<title>という単語列クラスに「天気—予報」が定義されているとすると、クラス連鎖モデル生成部311において、
25 形態素単位に分割された文表現は、「<SS>—明日—の—<title>—を—録画—して—<SE>」と置き換えられる。さらに、通常の単語

についてもクラスへの書き換えが行われて、「<SS>—<名詞>—<格助詞>—<title>—<格助詞>—<サ変名詞>—<動詞>—<SE>」などに置き換えられる。クラス連鎖モデル生成部 3 1 1 は、文表現コーパスからこのような系列を生成し、クラス C 1 の後にクラス C 2
5 の続く確率などを求めることのできるクラス連鎖モデルを生成する。

単語出力モデル生成部 3 1 2 においては、文表現コーパスの形態素解析結果の単語系列と、クラスシンボルに置き換えたクラス系列を元に、単語列クラスを除く単語について、クラスの出現回数とそれに該当する単語の出現回数の統計を取る。例えば、上記例であれば、<格助詞>の
10 出現回数は 2 回であり、このクラスに属する具体的な単語の出現回数は、「の」が 1 回、「を」が 1 回といったカウントを行う。この結果から、クラス C において単語 W が出現する確率などを求めることのできる、単語出力モデルが生成される。

なお、上記例においては、一般の単語の属するクラスとして文法的な
15 知識に基づくクラスを用いる場合を示したが、統計的に自動分類されたクラスを用いることも有効である。また、クラス連鎖モデルの例では、1 つ前を先行するクラスを条件とする確率モデルの例を示したが、2 つ前、3 つ前のクラスを条件とする確率モデルとすることも可能である。

次に、クラス依存単語 N グラムの生成について、説明する。

20 図 1 9 は、クラス依存単語 N グラム生成蓄積部 3 2 の内部機能構成を示すブロック図である。なお、図 6 と同一の番号を付されたブロックは実施の形態 1 において図 6 と共に示したものと同様であり、これらについては説明を割愛する。

図 1 9 に示されるように、クラス依存単語 N グラム生成蓄積部 3 2 は、
25 クラスコーパス蓄積部 1 2 1 や、クラス用形態素解析部 1 2 2、単語列クラス定義生成部 1 2 5、単語列クラス定義蓄積部 1 2 6 の他、単語列

クラスの除外単語をモデル作成の際に判定するモデル作成用除外単語判定部 3 2 1 と、クラス依存単語 N グラムを生成するクラス依存単語 N グラム生成部 3 2 2 とから構成される。

- 5 クラス依存単語 N グラム生成蓄積部 3 2 における処理は、実施の形態 1 に示したものと同様、まず、クラス用形態素解析部 1 2 2 において、クラスコーパス蓄積部 1 2 1 が蓄積している単語列を形態素解析し、単語に切り分ける。そして単語列クラス定義生成部 1 2 5 において、切り分けられた単語をもとに単語列クラス定義を生成して、単語列クラス定義蓄積部 1 2 6 に格納する。同時に、モデル作成用除外単語判定部 3 2
- 10 1 において、形態素に解析された単語について、単語のまま用いるか、除外単語とするかを判定する。そして、モデル作成用除外単語判定部 3 2 1 において除外単語と判定した場合、その除外単語の置き換えと、その除外単語を発音に対応するモーラに切り分ける置き換えとが行われる。

- 例えば「知床一の神秘」なる単語列において、モデル作成用除外単語判定部 3 2 1 における除外条件が「地名」であった場合、「<地名>一の神秘」と書き換えられる一方、「<MS>—シーレート—コー<ME>」とモーラの列に書き換えられる。ただし<MS>、<ME>は、除外単語のモーラ列の開始と終了を示す、仮想的なシンボルである。
- 15

- なお、「モーラ」（ここでは、日本語のモーラ、英語においては類似する音響単位として「シラブル」が存在する。）とは、日本語で 1 音（1 拍）と考えられている音韻のことであり、日本語をひらかな表記した時のひらかな 1 文字 1 文字に概ね対応する。また、俳句の 5・7・5 をカウントする時の 1 音に対応する。ただし、拗音（小さいやゆよの付く音）、促音（小さいっ/つまった音）、撥音（ん）については、1 音（1 拍）として発音されるか否かによって、独立した 1 つのモーラとして取り扱われたり、そうでなかったりする。例えば、「東京」であれば、4 つのモーラ
- 20
- 25

「と」、「う」、「きょ」、「う」から構成され、「札幌」であれば、4つのモーラ「さ」、「っ」、「ぼ」、「ろ」から構成され、「群馬」であれば、3つのモーラ「ぐ」、「ん」、「ま」から構成される。

5 クラス依存単語Nグラム生成部322は、除外単語部分を別の仮想的なシンボルに置き換えた単語の系列を、クラスコーパス中の多数のデータから求め、これを元に、それら単語列クラスにおける単語連鎖の頻度を確率モデル化し、クラス依存単語Nグラムを生成する。クラス依存単語Nグラム蓄積部124は、このクラス依存単語Nグラムが蓄積され、単語列仮説生成部80において、単語列クラスの生起確率の計算のために参照される。なお本実施の形態では、クラス依存単語Nグラムは単語列クラス内の単語の連鎖確率をモデル化するものとして示したが、前記のクラスNグラムのモデル生成でも示したように、一旦、単語を単語の属するクラスに置き換えて、クラスの連鎖確率とクラスー単語の出力確率の2種類の確率によってモデル化してもよい。

15 次に、クラス依存モーラNグラムの生成について説明する。

図20は、クラス依存モーラNグラム生成蓄積部33の内部機能構成を示すブロック図である。

図20に示されるように、クラス依存モーラNグラム生成蓄積部33は、

20 クラス依存単語Nグラム生成蓄積部32のモデル作成用除外単語判定部321から出力された除外単語の読みに相当するモーラの系列に基づいて、かかるモーラ列から除外単語におけるモーラの連鎖をモデル化するクラス依存モーラNグラム生成部331と、生成されたクラス依存モーラNグラムを蓄積するクラス依存モーラNグラム蓄積部332とから構成される。

25 クラス依存モーラNグラム生成蓄積部33における処理は、まず、ク

ラス依存単語 N グラム生成蓄積部 3 2 におけるモデル作成用除外単語判定部 3 2 1 において、除外単語と判定された単語の読みに相当するモーラの系列（例えば、「<MS>ーシーレートーコー<ME>」）がクラス依存モーラ N グラム生成部 3 3 1 に入力されると、クラス依存モーラ N グラム生成部 3 3 1 では、モーラ列に書き換えられたコーパス中の多数の除外単語が入力されて、モーラの連鎖の統計量が計算され、モーラ連鎖確率を示すモデルが生成される。具体的には、クラス依存モーラ N グラム生成部 3 3 1 は、単語 2 グラムモデルの場合、モーラについて、<MS>ーシ、シーレ、レート、トーコ、コー<ME>、…の頻度をそれぞれ求め、(M 1, M 2 の連鎖頻度) / (M 1 の頻度) を計算することにより、それぞれの確率 $P(M 2 | M 1)$ を求める。ただし、M 1, M 2 は、それぞれのモーラを示す。そして、クラス依存モーラ N グラム生成部 3 3 1 は、図 2 1 に示されるようにモーラ連鎖とその確率を対応付けてクラス依存モーラ N グラム蓄積部 3 3 2 に蓄積する。

15 クラス依存モーラ N グラム蓄積部 3 3 2 では、生成されたクラス依存モーラ N グラムが蓄積され、単語列仮説生成部 8 0 において、単語列クラスの生起確率の計算のために参照される。

なお、本実施の形態 3 では、クラス依存単語 N グラムの生成とクラス依存モーラ N グラムの生成に、実施の形態 1 と同一のクラスコーパス蓄積部 1 2 1 を利用するものとしたが、両者のモデルの生成のために、それぞれ異なるコーパスを用いてモデルを生成することも可能である。

認識処理における動作は、音声認識装置 1 の場合と同様に、音声入力に対して単語照合部 5 0 において単語照合がなされて単語仮説が生成され、単語列仮説生成部 8 0 において、単語区間を考慮した単語候補の接続と、単語系列に基づいた言語尤度の加算が行われて、単語列候補に対するスコアが計算される。ここにおいて、特定の単語列クラスに属する

単語列に対しては、認識用除外単語判定部 90 における除外単語の判断がなされ、クラス依存単語 N グラム生成蓄積部 32 に蓄積されたクラス依存単語 N グラム又は、クラス依存モーラ N グラム生成蓄積部 33 に蓄積されたクラス依存モーラ N グラムが参照されて言語尤度の計算が行われる。

ここで、本実施の形態 3 における言語尤度の計算方法を述べる。

単語列クラスを含む任意の単語列 $W_1, W_2, \dots, W_i, \dots, W_n$ に対して、それぞれの単語の属するクラスを、 $C_1, C_2, \dots, C_u, \dots, C_m$ とする。ただし、クラス C は単語列クラスである場合も含むものとする。単語列 $W_1 \sim W_n$ の中に単語列クラスに対応する系列が含まれており、これが部分列 W_j, \dots, W_k に対応するものとする。この時、単語列 $W_1 \sim W_n$ の言語尤度は、下記に示される式 (13) により計算される。

$$\begin{aligned} & \log P(W_1 \dots W_n) \\ &= \begin{cases} \sum_{i=1}^L \log P(C_u | C_{u-1}) \cdot P(W_i | C_u) & (C_u \text{ が単語列クラス以外の場合}) \\ \sum_{i=1}^L \log P(C_u | C_{u-1}) \cdot P_c(W_j \dots W_k | C_u) & (C_u \text{ が単語列クラスの場合}) \end{cases} \end{aligned} \quad \dots(13)$$

ただし、 $P(C_u | C_{u-1})$ および $P(W_i | C_u)$ は、クラス N グラムにより計算される確率である。 $P_c()$ は単語列クラスの生起確率であり、下記に示される式 (14) により計算される。

$$\begin{aligned} \log P_c(W_f \cdots W_k | C_u) \\ = \begin{cases} \sum_{i=1}^L \log P(W_i | W_{i-1}, C_u) & (W_i \text{が除外単語でない場合}) \\ \sum_{i=1}^L \log P_m(M_a \cdots M_b | \langle \text{除外単語} \rangle, C_u) \cdot P(\langle \text{除外単語} \rangle | W_{i-1}, C_u) & (W_i \text{が除外単語の場合}) \end{cases} \end{aligned} \quad \cdots(14)$$

ただし、 $P(W_i | W_{i-1}, C_u)$ および $P(\langle \text{除外単語} \rangle | W_{i-1}, C_u)$ は、クラス依存単語 N グラムによって計算される確率である。

- 5 また、 $M_a \cdots M_b$ は W_i の発音に対応するモーラ列を示しており、 $P_m()$ はクラス依存モーラ N グラムにより計算される確率である。

認識用除外単語判定部 90 は、上記確率計算式 (14) を第 1 形式において行うか、第 2 形式において行うかを判断する。この場合の判断基準は単語列クラス C_u および当該単語 W_i の言語的属性などの情報を用
10 いる。ここで言う言語的属性とは、例えば W_i が地名を示す固有名詞であるか否かといったものをさす。このように除外単語の判定を行って、除外単語であると判定されたものについては、モーラのごとく単語よりも短い単位に切り分けることで単語列クラスを表現するために、全ての単語を辞書に登録しておく必要がなく、またクラス依存単語 N グラムも、
15 あらゆる単語についての連鎖を考慮する必要がなくなるので、単語列クラスのコンパクトで高性能なモデル化が可能となる。

次に、具体的な例によって上記言語尤度の計算の例を示す。

例えば、単語列クラスとしてのタイトルクラスが含まれる例として、「太陽を一撃を一見」では、次のように言語尤度が、下記に示
20 される式 (15) により計算される。

$$\begin{aligned}
& \log P(<SS>, \text{太陽}, \text{を}, \text{撃て}, \text{を}, \text{見る}, <SE>) \\
&= \log P(<\text{タイトルクラス}> \mid <SS>) \cdot \\
&\quad P_c(<CS>, \text{太陽}, \text{を}, \text{撃て}, <CE> \mid <\text{タイトルクラス}>) \\
&+ \log P(<\text{格助詞}> \mid <\text{タイトルクラス}>) \cdot P(\text{を} \mid <\text{格助詞}>) \\
&+ \log P(<\text{動詞}> \mid <\text{格助詞}>) \cdot P(\text{見る} \mid <\text{動詞}>) \\
&+ \log P(<SE> \mid <\text{動詞}>) \cdot P(<SE> \mid <SE>)
\end{aligned}$$

…(15)

- ただし、＜ＳＳ＞および＜ＳＥ＞は、文頭および文末を意味する仮想的なシンボルである。また、＜ＣＳ＞および＜ＣＥ＞は単語列クラスの開始および終了を示す仮想的なシンボルである。ここで、タイトルクラス「太陽－を－撃て」の生起確率に基づく言語尤度は下記に示される式(16)によって計算される。

$$\begin{aligned}
& \log P_c(<CS>, \text{太陽}, \text{を}, \text{撃て}, <CE> \mid <\text{タイトルクラス}>) \\
&= \log P(\text{太陽} \mid <CS>, <\text{タイトルクラス}>) \\
&+ \log P(\text{を} \mid \text{太陽}, <\text{タイトルクラス}>) \\
&+ \log P(\text{撃て} \mid \text{を}, <\text{タイトルクラス}>) \\
&+ \log P(<CE> \mid \text{撃て}, <\text{タイトルクラス}>)
\end{aligned}$$

…(16)

- 上記例においては単語列クラスであるタイトルクラス中に除外単語が含まれていないと仮定しており、クラス依存モーラＮグラムは参照されていない。

次に単語列クラス中に除外単語が含まれる例として、「知床－の－神秘－を－見る」の場合について、言語尤度の計算法を下記に示される式(17)によって示す。

$$\begin{aligned}
& \log P(<SS>, \text{知床, の, 神秘, を, 見る, } <SE>) \\
&= \log P(<\text{タイトルクラス}> \mid <SS>) \cdot \\
&\quad P_c(<CS>, \text{知床, の, 神秘, } <CE> \mid <\text{タイトルクラス}>) \\
&+ \log P(<\text{格助詞}> \mid <\text{タイトルクラス}>) \cdot P(\text{を} \mid <\text{格助詞}>) \\
&+ \log P(<\text{動詞}> \mid <\text{格助詞}>) \cdot P(\text{見る} \mid <\text{動詞}>) \\
&+ \log P(<SE> \mid <\text{動詞}>) \cdot P(<SE> \mid <SE>) \\
&\dots(17)
\end{aligned}$$

ここで、タイトルクラスにおいて、地名を示す固有名詞は除外単語であるとする、「知床一の神秘」の生起確率に基づく言語尤度は下記に示される式（18）によって計算される。

$$\begin{aligned}
& \log P_c(<CS>, \text{知床, の, 神秘, } <CE> \mid <\text{タイトルクラス}>) \\
&= \log P_m(<MS>, \text{シ, レ, ト, コ, } <ME> \mid <\text{地名}>, <\text{タイトルクラス}>) \cdot \\
&\quad P(<\text{地名}> \mid <CS>, <\text{タイトルクラス}>) \\
&+ \log P(\text{の} \mid <\text{地名}>, <\text{タイトルクラス}>) \\
&+ \log P(\text{神秘} \mid \text{の}, <\text{タイトルクラス}>) \\
&+ \log P(<CE> \mid \text{神秘}, <\text{タイトルクラス}>) \\
&\dots(18)
\end{aligned}$$

5

ただし、<MS>および<ME>は、除外単語におけるモーラ系列の開始および終了を意味する仮想的なシンボルである。さら除外単語の生起確率 $P_m()$ は、クラス依存モーラNグラムに基づいて、言語尤度が下記に示される式（19）によって計算される。

$$\begin{aligned}
& \log P_m(<MS>, \text{シ, レ, ト, コ, } <ME> \mid <\text{地名}>, <\text{タイトルクラス}>) \\
&= \log P(\text{シ} \mid <MS>, <\text{地名}>, <\text{タイトルクラス}>) \\
&+ \log P(\text{レ} \mid \text{シ}, <\text{地名}>, <\text{タイトルクラス}>) \\
&+ \log P(\text{ト} \mid \text{レ}, <\text{地名}>, <\text{タイトルクラス}>) \\
&+ \log P(\text{コ} \mid \text{ト}, <\text{地名}>, <\text{タイトルクラス}>) \\
&+ \log P(<ME> \mid \text{コ}, <\text{地名}>, <\text{タイトルクラス}>) \\
&\dots(19)
\end{aligned}$$

10

つまり、「知床一の神秘を一見る」の場合、図22に示されるよう

に、＜SS＞－＜t i t l e＞－を－見る－＜SE＞の単語列について尤度が求められる。そして、＜t i t l e＞については、除外単語－の－神秘の単語列について尤度が求められる。さらに、除外単語については、＜MS＞－シーレート－コー＜ME＞のモーラ列について尤度が求められる。このように言語尤度を計算することで、タイトルクラスに含まれる全ての地名に基づいてクラス依存単語Nグラムを構築しなくとも、タイトルに含まれる地名を認識することが可能となる。

なお、本実施の形態3では、クラスNグラム、クラス依存単語Nグラム、クラス依存モーラNグラムの全てにおいて、直前の単語(モーラ)から当該単語(モーラ)へ接続する確率を用いる例を示したが、より長い履歴(例えば、2つ前まで、3つ前までなど)を考慮した確率モデルであってもよい。また上記言語モデルの言語単位は、単語(モーラ)を用いた例を示したが、連結単語(連結モーラ)を言語単位に含むモデルであってもよい。

また、単語クラスの例として、ここではタイトルクラスを示したが、この他、「総務省行政管理局」といった組織名クラス、「東名高速道路 海老名サービスエリア」といった施設名クラスなどを用いることも可能である。

さらに、単語列クラス中の除外単語として上記例では、「知床」のような地名を示したが、その他、「一郎」のような人名、「塩爺」のような流行語・新語、その他、種類数が多いなどの理由により、認識辞書に登録されていない単語や、統計的に除外単語としてモデル化することの効果が高いと判断された単語を対象とすることも有効である。

最後に、認識用除外単語判定部90について述べる。

認識用除外単語判定部90は、単語列クラスの生起確率に基づく言語尤度の計算において、クラス依存単語Nグラムのみによる計算を行うか、

クラス依存モーラNグラムも参照して計算を行うかを判定するものである。認識用除外単語判定部90での判定規則は、各種言語モデルの生成と同様、事前に決定されているべきものである。判定規則の例としては、
5 であるか否かといった規則が挙げられる。これについては、前述した通り、「一郎」のような人名、「塩爺」のような流行語・新語、その他、種類数が多いなどの理由により、認識辞書に登録されていない単語や、統計的に除外単語としてモデル化することの効果が高いと判断された単語を対象とすることも有効である。また、クラス依存単語Nグラム生成蓄積部32に含まれるモデル作成用除外単語判定部321は、前記の認識用除外単語判定部90と同一か又は同一の規則を蓄えるブロックであることが望ましい。しかし、言語モデルの構築時と認識時で意図的に除外単語の判定規則を異なるものにするという利用法も可能である。

なお、本実施の形態では、クラスNグラム言語モデル、クラス依存単語Nグラム言語モデル、クラス依存モーラNグラムの構築に、テキストコーパスを形態素解析した結果を利用する例を示したが、実施の形態2
15 に示したように、テキストコーパスを形態素解析し、さらに構文解析した結果から、クラスNグラム構築用の構文木と、クラス依存単語Nグラム構築用の構文木を求め、さらに構文的な条件を加味して除外単語判定
20 ルールを設けて、これを元に、クラス依存モーラNグラム構築用のデータを抽出して、各々言語モデルを構築することも有効である。この場合、構文木の上位層の構文木が構文木用クラスNグラム生成蓄積部に入力され、下位層の構文木のうち、除外単語判定部によって除外単語と判定された単語を、判定理由に関連する仮想シンボルに置き換えた構文木が、
25 構文木用クラス依存単語Nグラム生成構築部に入力され、除外単語判定部によって除外された単語が判定理由ごとにまとめられて、構文木用ク

ラス依存モーラ N グラム生成蓄積部に入力されて、夫々の言語モデルが構築される。

(実施の形態 4)

次に、本発明の実施の形態 4 について説明する。本実施の形態 4 による音声認識装置は、

実施の形態 1 (図 2) に示した、クラス N グラム生成蓄積部に蓄積されたクラス N グラムおよびクラス依存単語 N グラム生成蓄積部に蓄積されたクラス依存単語 N グラムを用いて、単語仮説を生成し認識結果を出力する点では、全く同様である。実施の形態 1 との相違点は、クラス依存単語 N グラム生成蓄積部がクラスコーパスの動的な変更に対応可能となっている点にある。

本実施の形態 4 におけるクラス依存単語 N グラム生成蓄積部の構成図を図 2 3 に示す。その他の図 6 と同じ番号を付与したブロックは、実施の形態 1 に示した処理と同じ処理を行うものとする。

図 2 3 に示されるように、クラス依存単語 N グラム生成蓄積部 1 3 は、電話回線やインターネットなどの通信手段によってクラス依存単語 N グラム構築に必要なコーパスを取得するクラスコーパス取得部 1 3 1 をさらに備える。

次に、本実施の形態 4 におけるクラス依存単語 N グラムの構築について説明する。

クラス依存単語 N グラム生成蓄積部は、予め定められた時間間隔に基づくトリガや、ユーザ操作に基づくトリガなどのトリガ信号によって、クラスコーパス取得部 1 3 1 によってクラスコーパスを取得する。取得されたクラスコーパスは実施の形態 1 に示したものと同様に、クラス用形態素解析部 1 2 2 およびクラス依存単語 N グラム生成部 1 2 3 によって、クラス依存単語 N グラムが生成される。

このように、クラス依存 N グラムを動的に更新できるようにすることの効果は、本実施の形態による音声認識装置を、例えばテレビ番組案内システムに用いた場合に顕著である。例えばテレビ番組案内システムへのユーザの発話、「明日一の太陽を撃てを一録画して」という文に
5 対して、クラス N グラムモデルが「明日一の<番組名>を一録画して」という言い回しをモデル化しており、クラス依存単語 N グラムが「太陽を一撃て」という番組名クラスをモデル化しているとする。この時、言い回し文型自体は時間による変化が少ないのに対して、番組名は日々放映される番組が変化するため大きく変化する。よって、必要に
10 応じて番組名コーパスを再取得し、番組名に対応するクラス依存単語 N グラムを再構築することにより、番組名を認識するためのモデルが最新のものに最適化されることになる。一方、言い回し文型に対応するクラス N グラムは、時間による変化が少ないため更新する必要はなく、事前にオフラインで構築しておいたクラス N グラムを蓄積するだけでよいた
15 め、計算リソースとハードウェアリソースの削減を可能とする。

なお本実施の形態 4 では、効果を示す例としてテレビ番組案内システムへの適用を示したが、その他、Web サイトの案内システム、図書の案内システム、カーナビゲーションシステムなどへの適用も効果的である。

また、本実施の形態では、下位の N グラムモデルである、クラス依存
20 単語 N グラム言語モデルのみを更新する例を示したが、上位 N グラム言語モデルのみを更新する、あるいは、上位と下位 N グラム言語モデルを夫々に応じたタイミングで更新するという方法も可能である。

さらに、本実施の形態では、クラス N グラム言語モデルおよびクラス依存単語 N グラム言語モデルを、夫々のモデル構築用のコーパスを用いて
25 オンラインで構築する例を示したが、別途オフラインで構築した夫々の言語モデルを、最適なタイミングで取得し、これを用いる方法も可能

である。

産業上の利用可能性

- 本発明は、機器への入力手段として音声認識技術を利用する種々の電子機器、例えばテレビ、ビデオなどのＡＶ機器、カーナビゲーションシステムなどの車載機器、ＰＤＡや携帯電話機などの携帯情報端末などに利用することができ、その産業用の利用可能性は非常に広くかつ大きい。
- 5

請 求 の 範 囲

1. 音声認識用の言語モデルを生成し、蓄積する言語モデル生成蓄積
5 装置であって、

複数のテキストを、特定の言語的特性を有する単語列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデルを生成し、蓄積する上位Nグラム言語モデル生成蓄積手段と、

前記単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルを生成し、蓄積する下位Nグラム言語モデル生成蓄積手段と
10 を備えることを特徴とする言語モデル生成蓄積装置。

2. 前記上位Nグラム言語モデル生成蓄積手段と、前記下位Nグラム言語モデル生成蓄積手段とは、異なるコーパスを用いて言語モデルをそれぞれ生成する
15

ことを特徴とする請求の範囲第1項記載の言語モデル生成蓄積装置。

3. 前記下位Nグラム言語モデル生成蓄積手段は、下位Nグラム言語モデル用のコーパスを更新するコーパス更新手段を有し、
20 下位Nグラム言語モデル生成蓄積手段は、更新されたコーパスに基づいて下位Nグラム言語モデルを更新生成する

ことを特徴とする請求の範囲第2項記載の言語モデル生成蓄積装置。

4. 前記下位Nグラム言語モデル生成蓄積手段は、前記単語列クラスの単語の系列を意味を有する最小の言語単位である1以上の形態素に解析し、当該単語列クラスに依存して当該形態素の系列をモデル化するこ
25

とにより前記下位 N グラム言語モデルを生成する

ことを特徴とする請求の範囲第 1 項記載の言語モデル生成蓄積装置。

- 5 5. 前記上位 N グラム言語モデル生成蓄積手段は、形態素に解析され
たテキストに含まれる前記単語列クラスを仮想的な仮想単語に置き換え、
当該仮想単語および各単語からなる系列をモデル化することにより前記
上位 N グラム言語モデルを生成する

ことを特徴とする請求の範囲第 1 項記載の言語モデル生成蓄積装置。

- 10 6. 前記下位 N グラム言語モデル生成蓄積手段は、前記単語列クラス
に出現する単語のうち、特定の単語を単語列クラスの構成単語に含めな
い除外単語とするか否かをその言語的特性に基づいて判断し、判断の結
果前記除外単語について単語の読みに対応するモーラおよびモーラを結
合した単位に切り分ける除外単語判断手段を有し、

- 15 前記言語モデル生成蓄積装置は、さらに

前記モーラおよびモーラを結合した単位の系列をモデル化して、確率
の対数値である言語尤度を前記単語列クラス又は該除外単語の言語的特
性に依存して付与することによりクラス依存モーラ N グラムを生成し、
蓄積するクラス依存モーラ N グラム生成蓄積手段

- 20 を備えることを特徴とする請求の範囲第 1 項記載の言語モデル生成蓄
積装置。

7. 前記言語モデル生成蓄積装置は、さらに、テキストを形態素解析
し、さらに構文解析し、予め決められた基準により選択された構文木上
25 のノードに着目して、前記テキストを複数の階層構造とする構文木を生
成する構文木生成手段を備え、

前記上位 N グラム言語モデル生成蓄積手段は、前記着目したノードから上層を構成する第 1 の部分木を用いて構文木用の前記上位 N グラム言語モデルを生成し、

5 前記下位 N グラム言語モデル生成蓄積手段は、前記着目したノードから下層を構成する第 2 の部分木を用いて生成された構文木用の下位 N グラム言語モデルを生成する

ことを特徴とする請求の範囲第 1 項記載の言語モデル生成蓄積装置。

8. 前記下位 N グラム言語モデル生成蓄積手段は、前記第 2 の層より
10 下層の部分木に出現する特定の単語を、予め決められた言語的特性に基づいて部分木の構成単語に含めない除外単語と判断する言語モデル生成用除外単語判断手段を有し、

前記下位 N グラム言語モデル生成蓄積手段は、前記除外単語については単語の読みに対応するモーラおよびモーラを結合した単位に分解し、
15 該モーラおよび結合モーラの系列を該除外単語の構文木上での位置および該除外単語の言語的特性に依存してモデル化することにより下位 N グラム言語モデルを生成する

ことを特徴とする請求の範囲第 7 項記載の言語モデル生成蓄積装置。

20 9. 前記言語モデル生成蓄積装置は、さらに、テキストを形態素解析し、さらに構文解析し、予め決められた基準により選択された構文木上のノードに着目して、前記テキストを複数の階層構造とする構文木を生成する構文木生成手段を備え、

前記上位 N グラム言語モデル生成蓄積手段は、前記構文木の最上位層
25 を構成する第 1 の部分木を用いて前記上位 N グラム言語モデルを生成し、

前記下位 N グラム言語モデル生成蓄積手段は、第 2 より下位の層を構

成する各部分木を、より上位の層における該部分木の位置付けに基づいて分類し、夫々の分類ごとの部分木を用いて前記下位 N グラム言語モデルを生成する

ことを特徴とする請求の範囲第 1 項記載の言語モデル生成蓄積装置。

5

10. 前記下位 N グラム言語モデル生成蓄積手段は、前記第 2 の層より下層の部分木に出現する特定の単語を、予め決められた言語的特性に基づいて部分木の構成単語に含めない除外単語と判断する言語モデル生成用除外単語判断手段を有し、

10 前記下位 N グラム言語モデル生成蓄積手段は、前記除外単語については単語の読みに対応するモーラおよびモーラを結合した単位に分解し、該モーラおよび結合モーラの系列を該除外単語の構文木上での位置および該除外単語の言語的特性に依存してモデル化することにより下位 N グラム言語モデルを生成する

15 ことを特徴とする請求の範囲第 9 項記載の言語モデル生成蓄積装置。

11. 前記上位 N グラム言語モデル生成蓄積手段は、前記単語列クラスを含む N 個の単語列とその単語の系列が生起する確率とを対応付けて上位 N グラム言語モデルを生成する

20 ことを特徴とする請求の範囲第 1 項記載の言語モデル生成蓄積装置。

12. 下位 N グラム言語モデル生成蓄積手段は、前記単語列クラスを構成する単語の長さ N の連鎖とその単語連鎖が生起する確率とを対応付けて下位 N グラム言語モデルを生成する

25 ことを特徴とする請求の範囲第 1 項記載の言語モデル生成蓄積装置。

13. 発声された単語の系列を認識する音声認識装置であって、

複数のテキストを、特定の言語的特性を有する単語列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデルと、

前記単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルと

を用いて音声を認識することを特徴とする音声認識装置。

14. 発声された単語の系列を認識する音声認識装置であって、

音声認識用の言語モデルを生成し、蓄積する言語モデル生成蓄積装置を備え、

前記言語モデル生成蓄積装置は、

複数のテキストを、特定の言語的特性を有する単語列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデルを生成し、蓄積する上位Nグラム言語モデル生成蓄積手段と、

15 前記単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルを生成し、蓄積する下位Nグラム言語モデル生成蓄積手段とを有し、

前記上位Nグラム言語モデル生成蓄積手段によって蓄積される上位Nグラム言語モデルと、前記下位Nグラム言語モデル生成蓄積手段によって蓄積される下位Nグラム言語モデルとを用いて音声を認識することを
20 特徴とする音声認識装置。

15. 前記上位Nグラム言語モデルと、前記下位Nグラム言語モデルとは、異なるコーパスを用いてそれぞれ生成され、

前記音声認識装置は、異なるコーパスを用いてそれぞれ構築された前
25 記上位Nグラム言語モデルと、前記下位Nグラム言語モデルとを用いて
音声を認識する

ことを特徴とする請求の範囲第 1 4 項記載の音声認識装置。

1 6 . 前記下位 N グラム言語モデル生成蓄積手段は、下位 N グラム言語モデル用のコーパスを更新するコーパス更新手段を有し、

5 下位 N グラム言語モデル生成蓄積手段は、更新されたコーパスに基づいて下位 N グラム言語モデルを更新生成し、

前記音声認識装置は、更新された前記下位 N グラム言語モデルを用いて音声を認識する

ことを特徴とする請求の範囲第 1 5 項記載の音声認識装置。

10

1 7 . 前記下位 N グラム言語モデル生成蓄積手段は、前記単語列クラスの単語の系列を意味を有する最小の言語単位である 1 以上の形態素に解析し、当該単語列クラスに依存して当該形態素の系列をモデル化することにより前記下位 N グラム言語モデルを生成し、

15 前記音声認識装置は、前記形態素の系列にモデル化された前記下位 N グラム言語モデルを用いて音声を認識する

ことを特徴とする請求の範囲第 1 4 項記載の音声認識装置。

20 1 8 . 前記上位 N グラム言語モデル生成蓄積手段は、形態素に解析されたテキストに含まれる前記単語列クラスを仮想的な仮想単語に置き換え、当該仮想単語および各単語からなる系列をモデル化することにより前記上位 N グラム言語モデルを生成し、

前記音声認識装置は、前記仮想単語および各単語からなる系列にモデル化された前記上位 N グラム言語モデルを用いて音声を認識する

25 ことを特徴とする請求の範囲第 1 4 項記載の音声認識装置。

19. 前記下位Nグラム言語モデル生成蓄積手段は、前記単語列クラスに出現する単語のうち、特定の単語を単語列クラスの構成単語に含めない除外単語とするか否かをその言語的特性に基づいて判断し、判断の結果前記除外単語について単語の読みに対応するモーラおよびモーラを結合した単位に切り分ける除外単語判断手段を有し、
- 前記言語モデル生成蓄積装置は、さらに
- 前記モーラおよびモーラを結合した単位の系列をモデル化して、確率の対数値である言語尤度を前記単語列クラス又は該除外単語の言語的特性に依存して付与することによりクラス依存モーラNグラムを生成し、
- 10 蓄積するクラス依存モーラNグラム生成蓄積手段を備え、
- 前記音声認識装置は、前記クラス依存モーラNグラムを用いて音声を認識する
- ことを特徴とする請求の範囲第18項記載の音声認識装置。
- 15 20. 前記言語モデル生成蓄積装置は、さらに、テキストを形態素解析し、さらに構文解析し、予め決められた基準により選択された構文木上のノードに着目して、前記テキストを複数の階層構造とする構文木を生成する構文木生成手段を備え、
- 前記上位Nグラム言語モデル生成蓄積手段は、前記着目したノードから上層を構成する第1の部分木を用いて構文木用の前記上位Nグラム言語モデルを生成し、
- 20 前記下位Nグラム言語モデル生成蓄積手段は、前記着目したノードから下層を構成する第2の部分木を用いて生成された構文木用の下位Nグラム言語モデルを生成し、
- 25 前記音声認識装置は、
- 音声から特徴パラメータを生成する音響処理手段と、

単語の発音および音響的特徴を照合し、単語の発声区間および単語の音響尤度を含む単語仮説群を生成する単語照合手段と、

構文木用の前記上位 N グラム言語モデルと構文木用の下位 N グラム言語モデルとを参照し、前記単語仮説群から単語列仮説を生成し、音声の

5 認識結果を生成する単語列仮説生成手段と

を備えることを特徴とする請求の範囲第 19 項記載の音声認識装置。

21. 前記下位 N グラム言語モデル生成蓄積手段は、前記第 2 の層より下層の部分木に出現する特定の単語を、予め決められた言語的特性に
10 基づいて部分木の構成単語に含めない除外単語と判断する言語モデル生成用除外単語判断手段を有し、

前記下位 N グラム言語モデル生成蓄積手段は、前記除外単語については単語の読みに対応するモーラおよびモーラを結合した単位に分解し、
該モーラおよび結合モーラの系列を該除外単語の構文木上での位置および
15 該除外単語の言語的特性に依存してモデル化することにより下位 N グラム言語モデルを生成し、

前記単語列仮説生成手段は、音声の認識結果を生成する

ことを特徴とする請求の範囲第 20 項記載の音声認識装置。

20 22. 前記言語モデル生成蓄積装置は、さらに、テキストを形態素解析し、さらに構文解析し、予め決められた基準により選択された構文木上のノードに着目して、前記テキストを複数の階層構造とする構文木を生成する構文木生成手段を備え、

前記上位 N グラム言語モデル生成蓄積手段は、前記構文木の最上位層
25 を構成する第 1 の部分木を用いて前記上位 N グラム言語モデルを生成し、

前記下位 N グラム言語モデル生成蓄積手段は、第 2 より下位の層を構

成する各部分木を、より上位の層における該部分木の位置付けに基づいて分類し、夫々の分類ごとの部分木を用いて前記下位 N グラム言語モデルを生成し、

前記音声認識装置は、第 1 の部分木を用いて生成された前記上位 N グラム言語モデルと第 2 より下位の層を構成する各部分木を用いて生成された前記下位 N グラム言語モデルとを用いて音声認識すること
5 ことを特徴とする請求の範囲第 1 4 項記載の音声認識装置。

2 3. 前記下位 N グラム言語モデル生成蓄積手段は、前記第 2 の層より下層の部分木に出現する特定の単語を、予め決められた言語的特性に基づいて部分木の構成単語に含めない除外単語と判断する言語モデル生成用除外単語判断手段を有し、

前記下位 N グラム言語モデル生成蓄積手段は、前記除外単語については単語の読みに対応するモーラおよびモーラを結合した単位に分解し、
15 当該モーラおよび結合モーラの系列を該除外単語の構文木上での位置および該除外単語の言語的特性に依存してモデル化することにより下位 N グラム言語モデルを生成し、

前記音声認識装置は、前記除外単語を含めない前記上位 N グラム言語モデルと前記除外単語を含めた前記下位 N グラム言語モデルとを用いて
20 音声を認識する

ことを特徴とする請求の範囲第 2 2 項記載の音声認識装置。

2 4. 前記上位 N グラム言語モデル生成蓄積手段は、前記単語列クラスを含む長さ N の単語連鎖とその単語連鎖が生起する確率とを対応付けて上位 N グラム言語モデルを生成し、
25

前記音声認識装置は、前記単語列クラスを含む N 個の単語列が生起す

る確率を乗算することにより単語列仮説を評価する単語列仮説生成手段を有する

ことを特徴とする請求の範囲第14項記載の音声認識装置。

- 5 25. 下位Nグラム言語モデル生成蓄積手段は、前記単語列クラスを構成する単語の長さNの単語連鎖とその単語連鎖が生起する確率とを対応付けて下位Nグラム言語モデルを生成し、

前記音声認識装置は、前記単語列クラス内のN個の単語列が生起する確率を乗算することにより単語列仮説を評価する単語列仮説生成手段を

- 10 有する

ことを特徴とする請求の範囲第14項記載の音声認識装置。

26. 音声認識用の言語モデルを生成する言語モデル生成方法であって、

- 15 複数のテキストを、特定の言語的特性を有する単語列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデルを生成する上位Nグラム言語モデル生成ステップと、

前記単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルを生成する下位Nグラム言語モデル生成ステップと

- 20 を含むことを特徴とする言語モデル生成方法。

27. 発声された単語の系列を認識する音声認識方法であって、

複数のテキストを、特定の言語的特性を有する単語列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデルと、前記単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルとを用いて
25 音声認識することを特徴とする音声認識方法。

28. 前記音声認識方法は、

単語列によって特定の言語的性質を有する単語列を単語列クラスとしてまとめ、該単語列クラスを、構成単語間の言語的な関係に基づき該単語列クラスに依存してモデル化したクラス依存単語Nグラムによって、

5 確率の対数値である言語尤度を付与するステップと、

テキストを単語および上記単語列クラスに解析し、該単語および単語列クラスの系列を言語的な関係に基づきモデル化したクラスNグラムによって、確率の対数値である言語尤度を付与するステップと、

10 連続した音声から抽出された特徴パラメータに、単語の発音および音響的特徴を照合し、単語の発声区間および単語の音響尤度を含む単語仮説群を生成し、前記クラスNグラムおよび前記クラス依存単語Nグラムを参照し、前記単語仮説群から単語列仮説を生成し、音声の認識結果を出力するステップと

を含むことを特徴とする音声認識方法。

15

29. 音声認識用の言語モデルを生成する言語モデル生成方法を実行するためのプログラムであって、

複数のテキストを、特定の言語的特性を有する単語列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデルを生成する上位N

20 グラム言語モデル生成ステップと、

前記単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルを生成する下位Nグラム言語モデル生成ステップと

をコンピュータに実行させることを特徴とするプログラム。

25 30. 発声された単語の系列を認識する音声認識方法を実行するためのプログラムであって、

複数のテキストを、特定の言語的特性を有する単語列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデルと、前記単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルとを用いて音声を認識するステップをコンピュータに実行させることを特徴とする

5 プログラム。

要 約 書

- 音声認識用の言語モデルを生成し、蓄積する言語モデル生成蓄積装置
- (10) は、複数のテキストを、特定の言語的特性を有する単語列クラスを含む単語の系列としてモデル化した上位Nグラム言語モデルを生成し、蓄積する上位のクラスNグラム生成蓄積部(11)と、単語列クラス内の単語の系列をモデル化した下位Nグラム言語モデルを生成し、蓄積する下位のクラス依存単語Nグラム生成蓄積部(12)とを備えることを特徴とする。
- 5

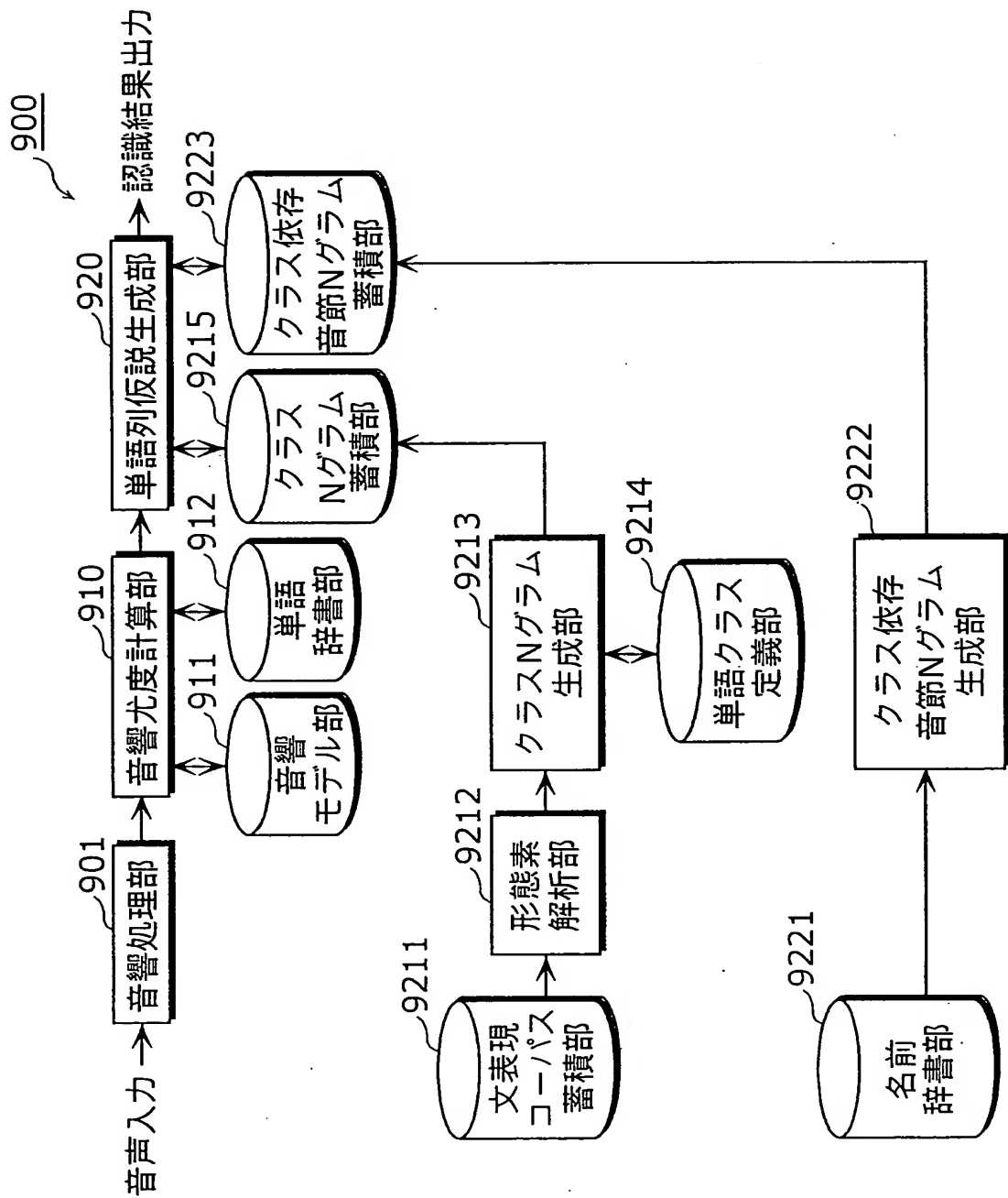


図1

図2

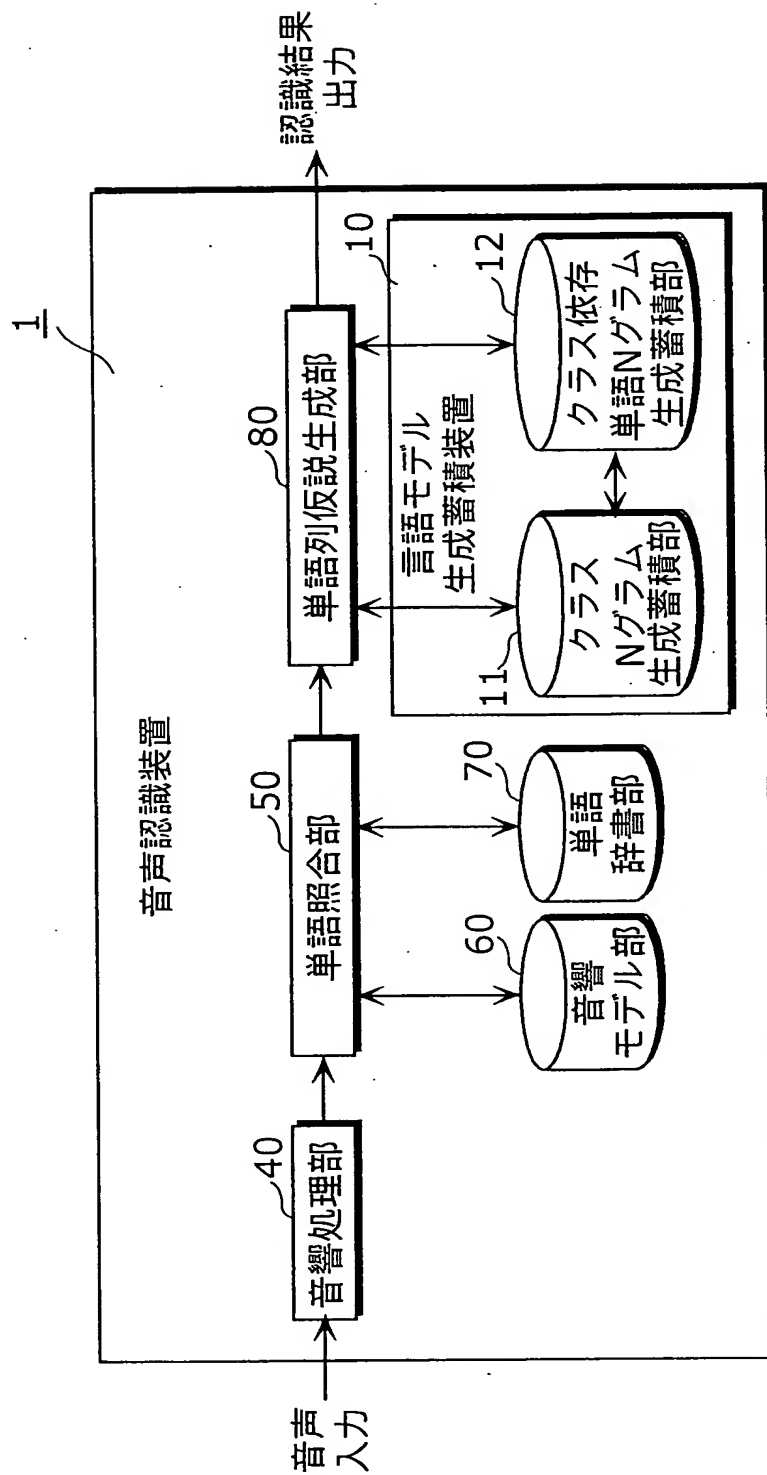


図3

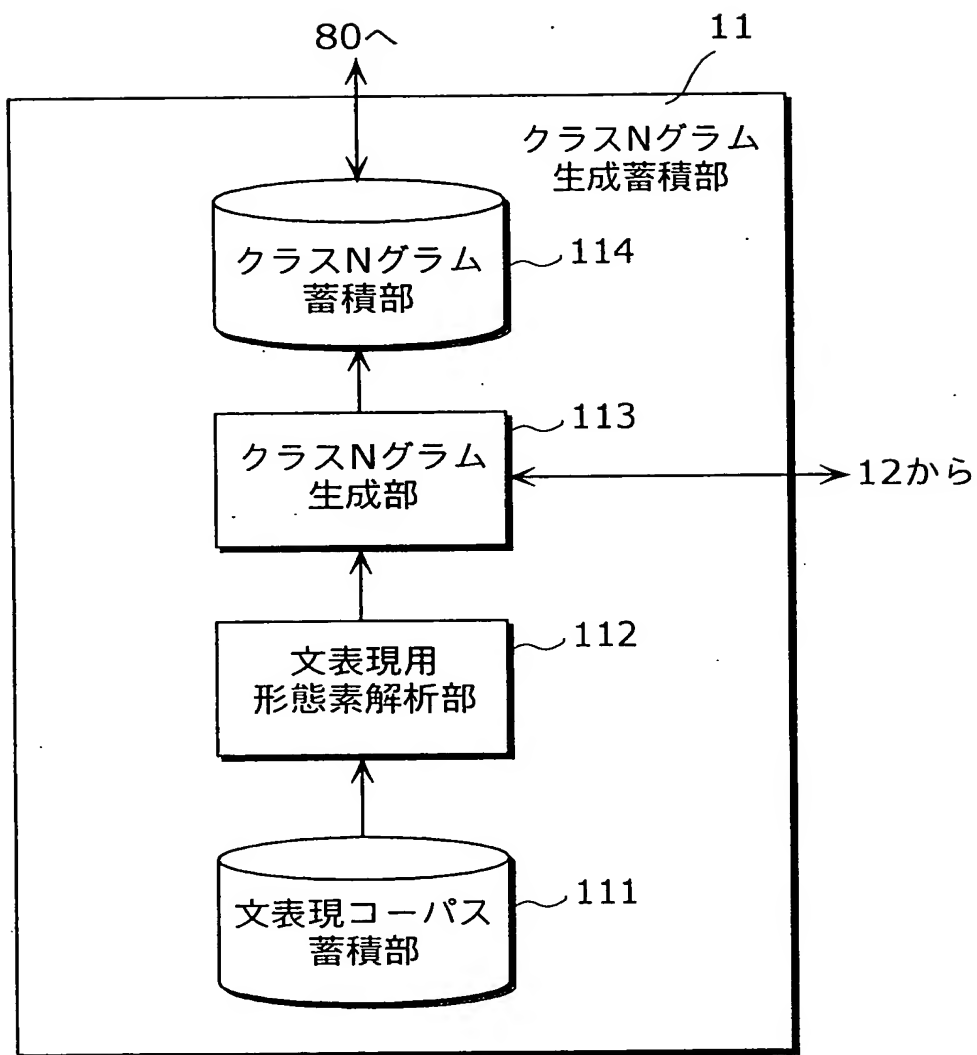
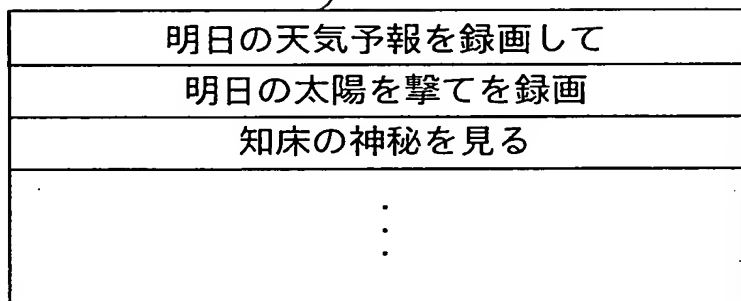


図4

文表現コーパス蓄積部111



明日の天気予報を録画して
明日の太陽を撃てを録画
知床の神秘を見る
⋮

図5

クラスNグラム蓄積部114

単語列	$P(W_i W_{i-1})$
<SS> - 明日	0.45
明日 - の	0.41
の - <title>	0.35
<title> - を	0.59
を - 録画	0.43
録画 - して	0.35
して - <SE>	0.33
録画 - <SE>	0.31
<SS> - <title>	0.51
⋮	⋮

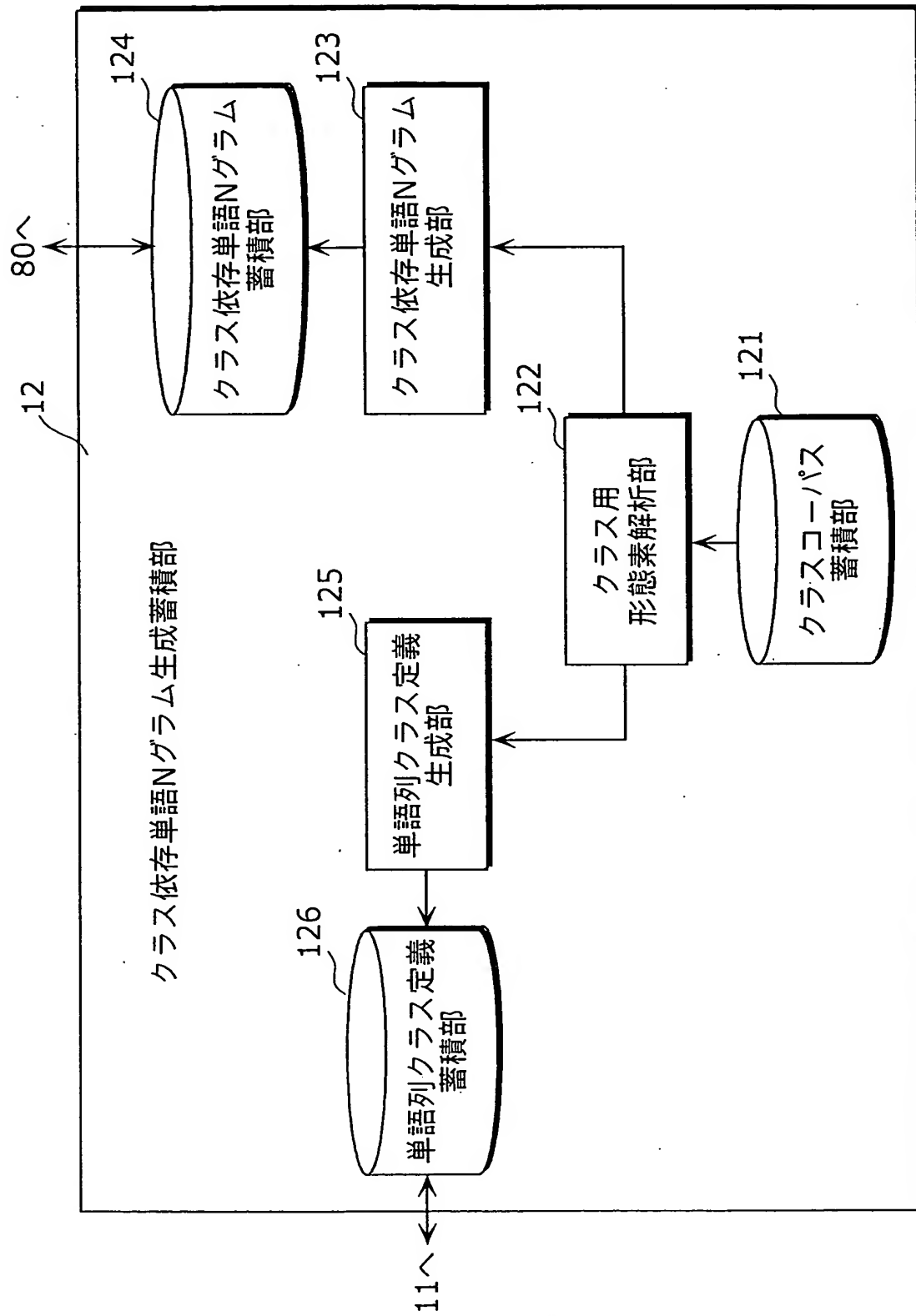



図6

図7

クラスコーパス蓄積部121



天気予報
太陽を撃て
知床の神秘
⋮
チャーリー海
池乃キングヨ
⋮

図8

クラス依存単語Nグラム蓄積部124

クラス単語列	$P(W_i W_{i-1})$
<CS>－天気	0.11
天気－予報	0.15
予報－<CE>	0.16
<CS>－太陽	0.13
太陽－を	0.11
を－撃て	0.10
撃て－<CE>	0.12
<CS>－知床	0.09
知床－の	0.08
の－神秘	0.07
神秘－<CE>	0.11
⋮	⋮
<CS>－池乃	0.08
池乃－キングョ	0.05
キングョ－<CE>	0.07
<CS>－チャーリー	0.07
チャーリー－海	0.05
海－<CE>	0.16
⋮	⋮

図9

単語列クラス定義蓄積部126

単語列クラス	単語列
title	<CS>－天気－予報－<CE>
	<CS>－太陽－を－撃て－<CE>
	<CS>－知床－の－神秘－<CE>
	⋮
人名	<CS>－チャーリー－海－<CE>
	<CS>－池乃－キングヨ－<CE>
	⋮
⋮	⋮

図10

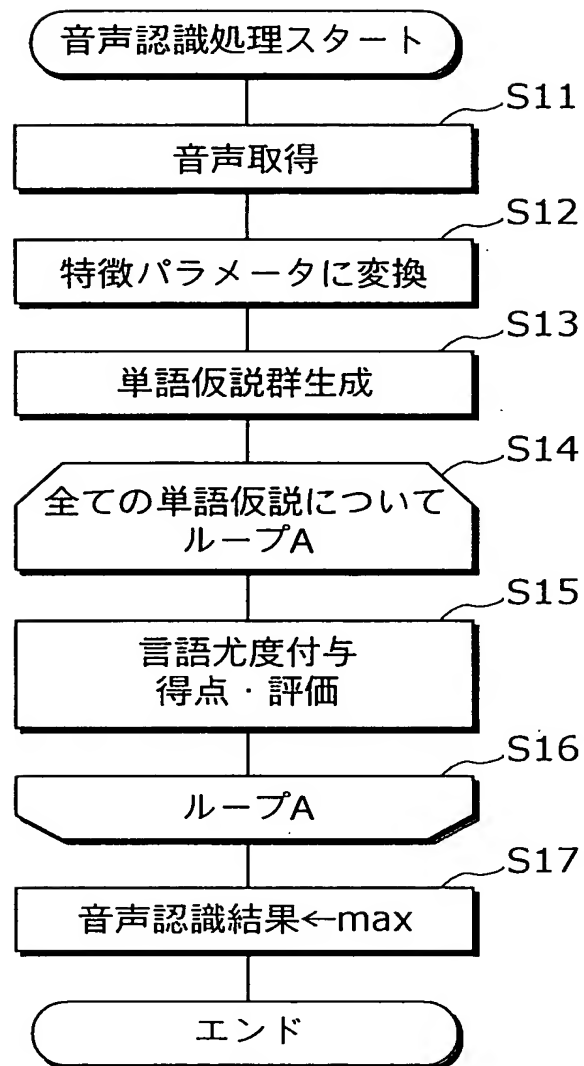


図11

単語列	確率
$\langle SS \rangle - a - b - c - d - \langle SE \rangle$	$P(a, b, c, d)$
$\langle SS \rangle - C - b - c - d - \langle SE \rangle$	$P(C, b, c, d) \cdot P(a \mid C)$
$\langle SS \rangle - C - c - d - \langle SE \rangle$	$P(C, c, d) \cdot P(a, b \mid C)$
\vdots	\vdots
$\langle SS \rangle - a - C - c - d - \langle SE \rangle$	$P(a, C, c, d) \cdot P(b \mid C)$
$\langle SS \rangle - a - C - d - \langle SE \rangle$	$P(a, C, d) \cdot P(b, c \mid C)$
\vdots	\vdots
\vdots	\vdots
\vdots	\vdots
$\langle SS \rangle - C - \langle SE \rangle$	$P(a, b, c, d \mid C)$

図12

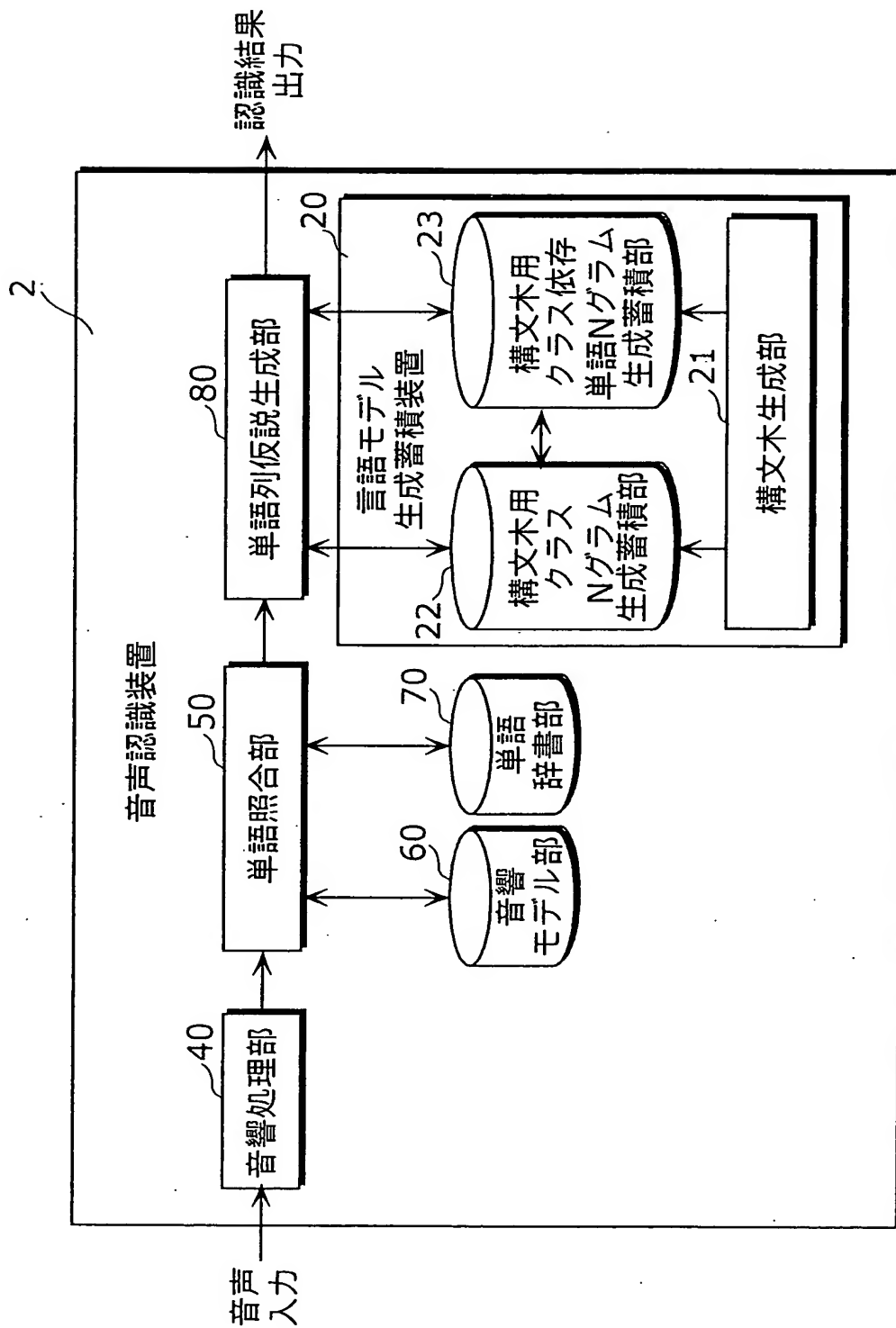


図13

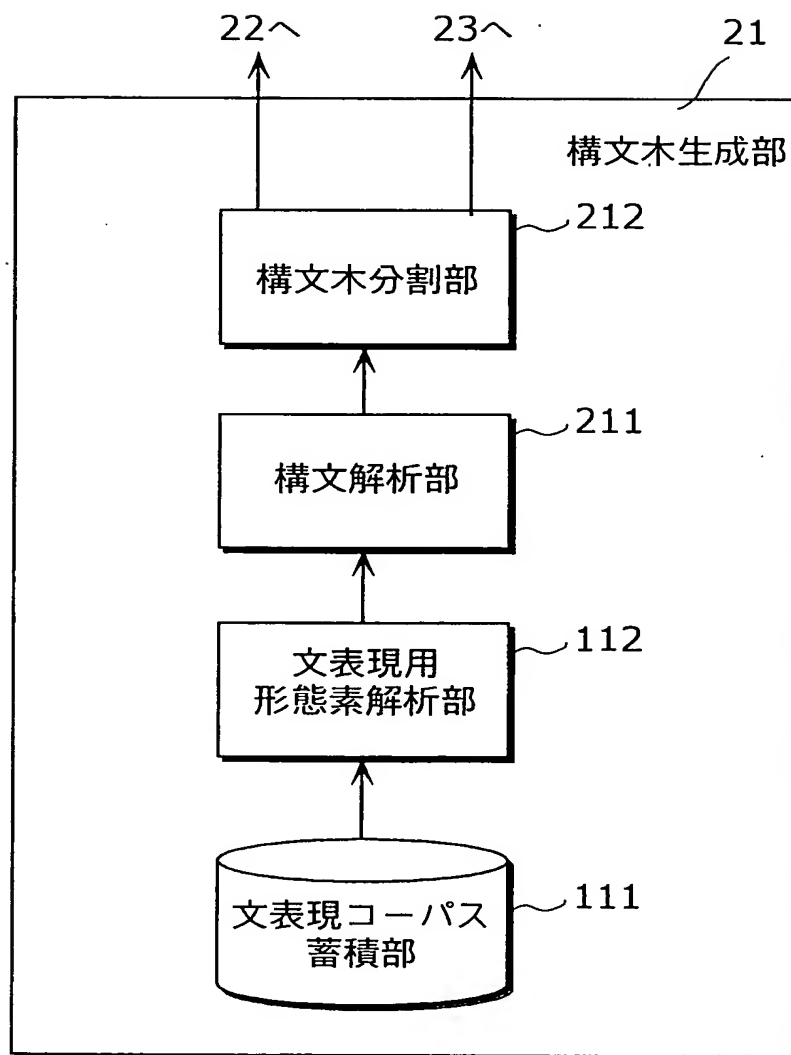


図14

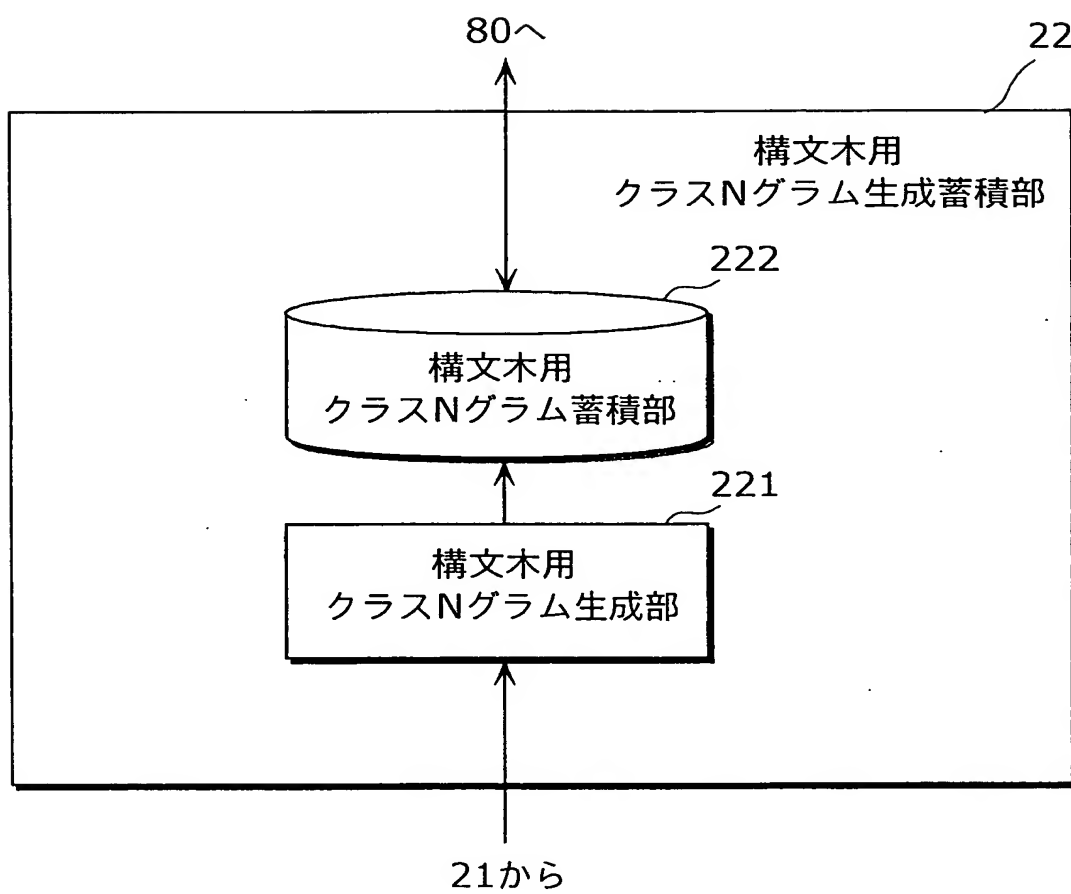


図15

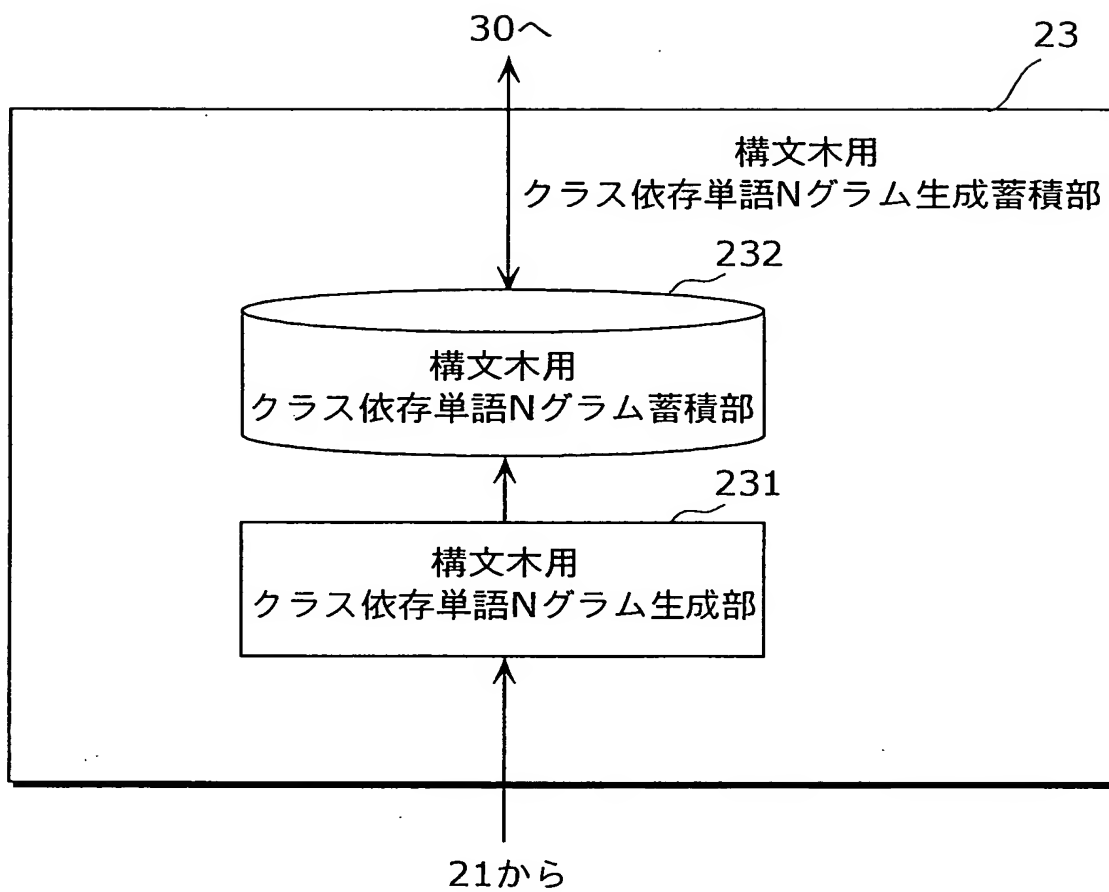


図16

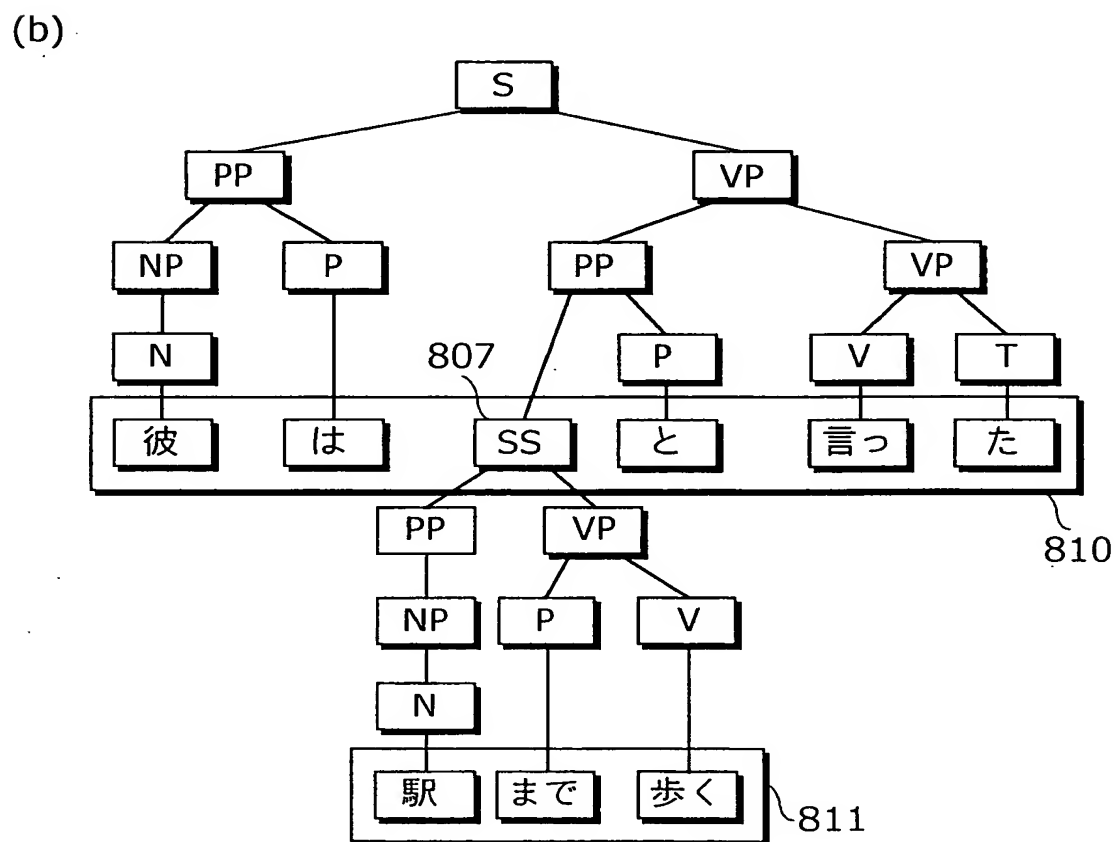
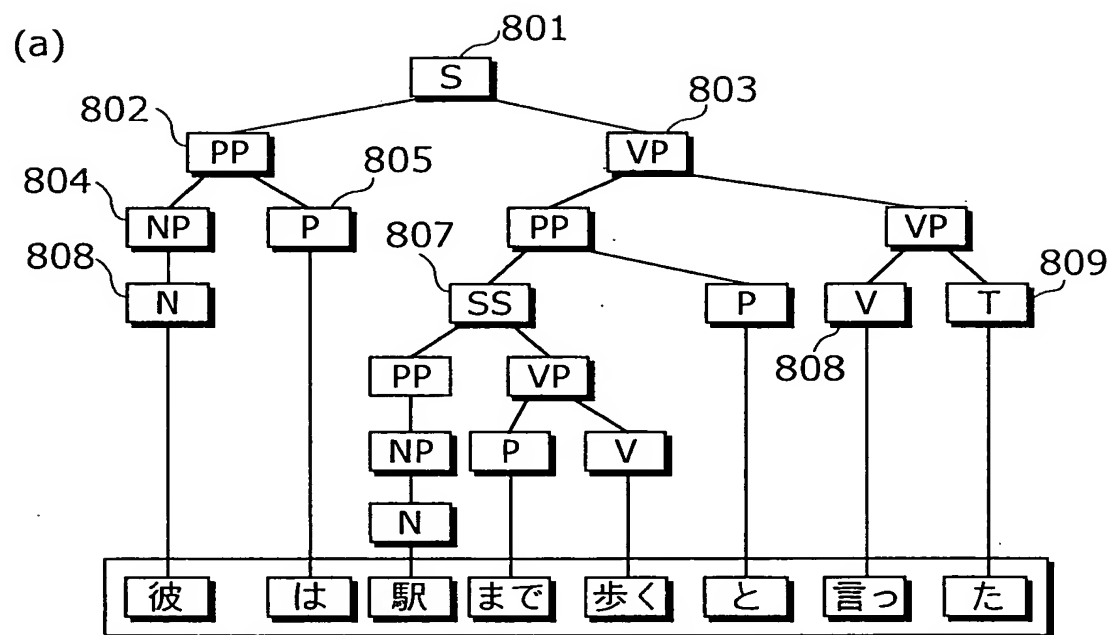


図17

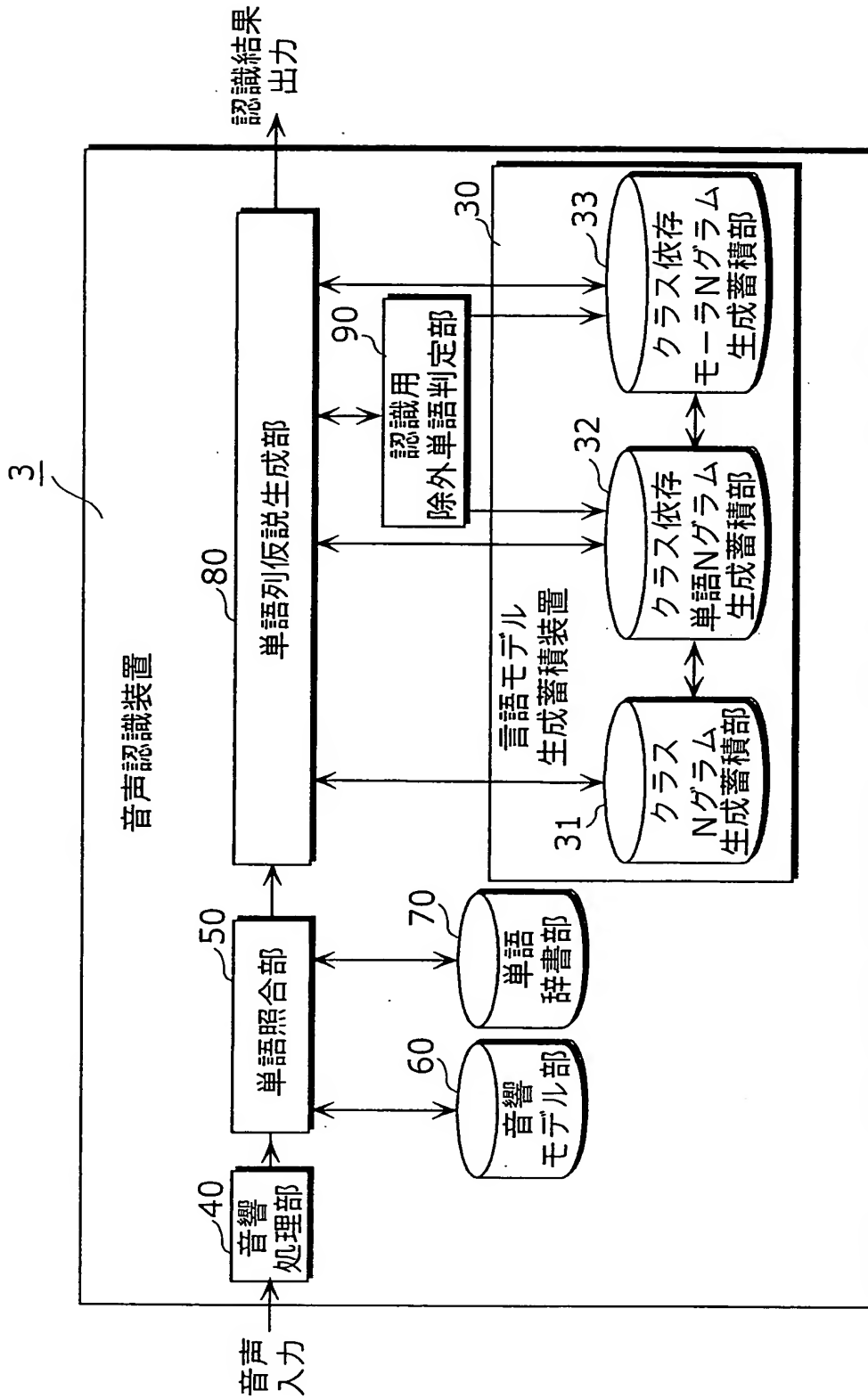


図18

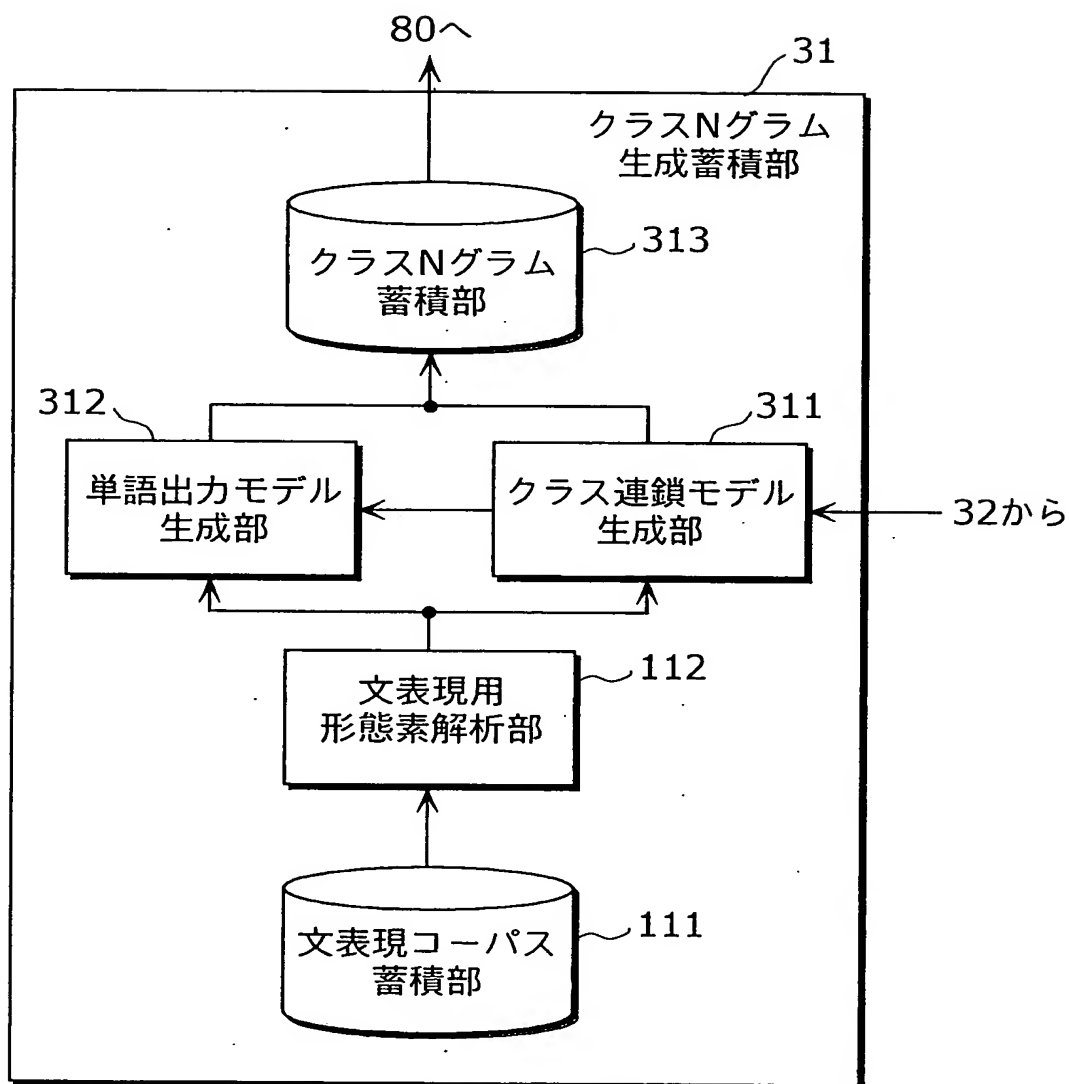


図19

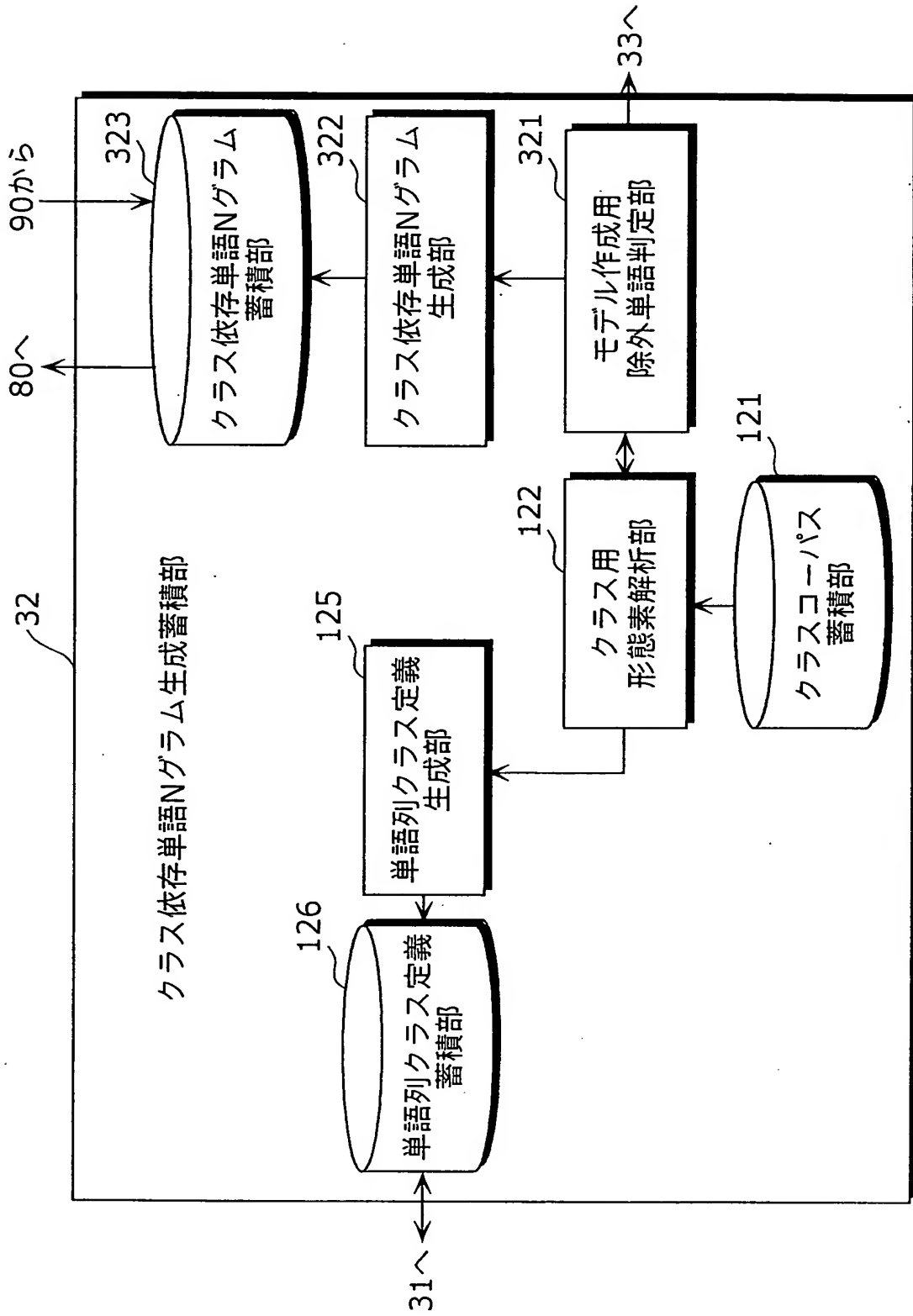


図20

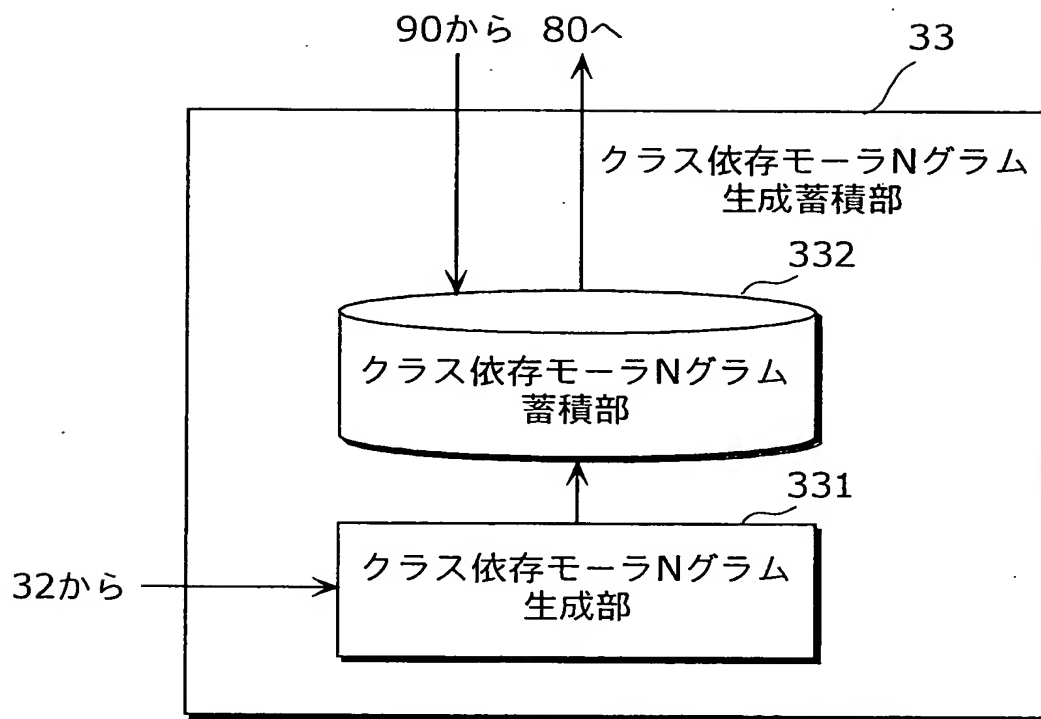


図21

クラス依存モーラNグラム蓄積部332

モーラ列	確率
<MS>-シ	0.10
シーレ	0.09
レート	0.06
ト-コ	0.07
コ- <ME>	0.08
⋮	⋮

図22

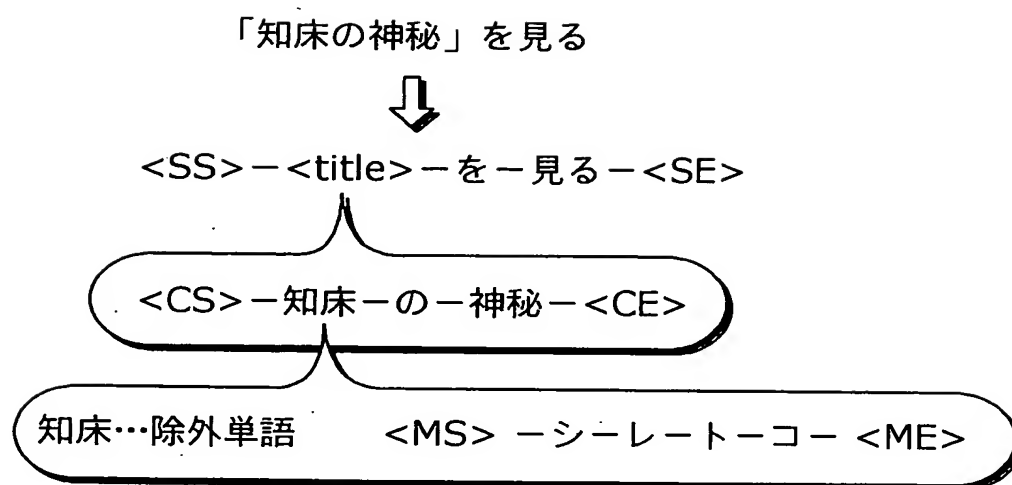


図23

